NEUROPSYCHOLOGICAL DIFFERENCES BETWEEN PHYSICIANS REFERRED FOR COMPETENCY EVALUATIONS AND A CONTROL GROUP OF PHYSICIANS

A Dissertation

Presented to

The College of Education

University of Denver

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

---

by

Lauri L. Korinek

June 2005

Advisor: Cynthia McRae, Ph.D.

GRADUATE STUDIES

AT

THE UNIVERSITY OF DENVER

Upon the recommendation of the Dean of the College of Education this dissertation is hereby accepted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Cynthia McRae

Professor in charge of dissertation

James R. Moran

Vice Provost for Graduate Studies

6/3/05

Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In 1973 the American Medical Association Council on Mental Health published an article on "The Sick Physician." This article presented the concern that many physicians are impaired, placing the public at risk of harm. The article suggested that these physicians needed to be identified, assessed, and treated. According to the American Medical Association Counsel article ("The Sick Physician"), an impaired physician is "one who is unable to practice medicine with reasonable skill and safety to patients because of physical or mental illness including deterioration through the aging process, or loss of motor skill, or excessive use of drugs including alcohol" (p. 684). The American Medical Association called on physicians to self-monitor and confront or identify fellow physicians who were impaired.

Studies suggest that about 7% to 10% of physicians practicing medicine in the United States are impaired in some fashion (Hall et al., 1999; Van Komen, 2000). In actual number that means that between 57,000 and 80,000 physicians practicing in the year 2000 were impaired (Pasko & Seidman, 2002). The etiology of impairment in physicians is multifaceted. Examples of physician behaviors that could place patients at risk include unethical conduct, prescribing

problems, documentation problems (Stulp deGroot, et al., 2004), inappropriate boundaries with patients, poor outcomes, or not staying current on treatments and medications. Until recently, cognitive factors that potentially impact physician functioning have been either overlooked or given only minimal attention as possibly contributing to physician impairment. Some factors related to cognitive functioning that could impact physician performance include normal aging, illnesses or injuries that adversely impact cognitive functioning, sustained abuse of drugs and alcohol, and some psychiatric problems such as severe depression.

Concerning etiologies that adversely impact cognitive functioning, the normal aging process "may create changes which make it increasingly hard for the physician to maintain the standard of practice" (Sadavoy, 1994, p. 265). As physicians age, cognitive decline is expected. More specifically, physicians can expect decline in motor response times, long-term concentration, the ability to identify essential versus nonessential details, visual-spatial ability, the ability to multibank, and the ability to learn new material in a short span of time (Goldstein, 2000; Powell & Whitla, 1994). A decline in different aspects of memory, including visual-spatial memory, also accompanies aging (Wilson et al., 2002; Zelinski & Kennison, 2001).

One factor that is important to consider when trying to understand the impact of normal aging on physician performance is that variability in cognitive

functioning is greater in older physicians (Powell & Whitla, 1994). Among younger physicians it can be assumed that cognitive functioning is generally quite high. However, as physicians age, it is clear that some physicians' cognitive functioning is maintained at very high levels, while other physicians' cognitive functioning declines significantly with age. Powell and Whitla found, in their physician population, that between 1/5 and 1/3 of the octogenarians maintained cognitive functioning (based on accuracy, not speed) at the same level as individuals in their prime age group (35 - 44 years of age). Thus, even though the overall mean of physician cognitive functioning declines with age, it appears a significant proportion of older physicians continue to function at high cognitive levels.

Along with the cognitive decline associated with normal aging there is also the potential for cognitive decline through illness or injury such as brain tumors, "cerebral vascular disease, head injury, Alzheimer's disease, demyelinating disease, Parkinson's disease, brain infections, and toxic and metabolic brain disorder" (Madden, 1988, p. 201). In a 7.5 year study of individuals from the age of 55 to 85+, 12.5% developed dementia, and another 10% were suspected to have early dementia (Kawas, Gray, Brookmeyer, Fozard, & Zonderman, 2000). As individuals in this study aged, the incidence rate for dementia increased, starting at .08% for individuals in the 60-64 age range and steadily increasing to an incidence rate of 6.48% in individuals 85 and older

(Kawas et al.). While the participants in this study were not necessarily physicians, they were well-educated with approximately half of the sample having 17-25 years of education, constituting a group comparable to physicians. Thus, it appears that even as well-educated professionals (including physicians) age, the chances for diseases that adversely impact cognitive functioning increase.

Intuitively, an understanding that physicians' cognitive abilities decline with age suggests that their clinical performance may also decline. Research bears this out, demonstrating that as physicians age their clinical performance (as a group) declines (Choudhry, Fletcher, & Soumeral, 2005; McAuley, Paul, Morrison, Beckett, & Goldsmith, 1990; Norcini, Lipner, Benson, & Webster, 1985; Norman et al., 1993). Older physicians tend to perform worse than younger physicians on tests of professional knowledge, review of patient charts, oral exams, and meetings with hypothetical patients (Norman et al.). Thus, some older physicians may present a threat to public safety and a threat to the public trust of the medical profession.

Another factor that can contribute to impairment is related to physicians' use of drugs and alcohol, with alcohol being the most frequently abused substance (Hughes et al., 1992). Estimates of alcohol abuse range from 7% to 13% in practicing physicians (McAuliffe et al., 1991; Moore, Mead, & Pearson, 1990; Regier et al., 1990).

The use of drugs by physicians also contributes to physician impairment. While physicians are less likely than the general public to use illicit drugs, they are more likely to abuse prescription medications (Hughes et al., 1992). The two most commonly abused medications are opiates and benzodiazepines. According to Hughes et al., approximately 4% of physicians admit to having a drug abuse or dependence problem at some point in their lifetime. While this study did not directly address the impact of current substance use on physician competency, long term drug and alcohol usage can result in permanent neurological impairment, which may adversely impact a physician's ability to practice medicine safely.

To address the problem of all impaired physicians, medical boards have the responsibility of identifying physicians at risk and intervening before the public is harmed (Clay & Conaster, 2003). Hospitals also review physician performance and have protocols to address substandard performance. As a result of a review by a medical board or hospital, more information about why a particular physician is performing poorly might be needed. Therefore, programs that are designed to provide thorough evaluations of physicians have been developed.

Initially, these evaluation programs did not include any form of neuropsychological testing or screening and many physicians have been resistant to the idea of using neuropsychological assessment in competency evaluations.

However, recent research conducted through physician evaluation programs suggests that neuropsychological concerns do impact the group of physicians evaluated (Turnbull et al., 2000; Williams, Williams, & Norcross, 2002). One such program, the Physician Review Program (PREP), was developed through the College of Physicians and Surgeons of Ontario. In this program, when physicians fail to demonstrate competency, they are referred for an intensive one-day evaluation. This evaluation incorporates several modalities of assessment such as "multiple choice questions; peer-observed encounters with standardized patients; chart-simulated recall; and structured office oral examinations" (Hanna, Premi, & Turnbull, 2000, p. 174). Physicians are then rated according to their level of clinical performance and individualized educational plans are developed.

A PREP research project was developed for several physicians who were unsuccessful in completing their educational plan (Hanna, Premi, & Turnbull, 2000). In the study, five physicians were selected who had failed a self-directed remedial continuing medical education (CME) program, and who were rated as having moderate to major problems. The competency of each person was seriously questioned. However, no neuropsychological testing was conducted with these physicians. These physicians went through an extensive and concentrated three year program to improve their clinical performance. At the conclusion of the program only one physician's performance improved, one

physician's performance remained the same, and three physicians' performance declined. In discussing the reasons for the lack of improvement in the majority of physicians in this program, the authors suggested that "it is possible that their incompetence arose from early age-related cognitive decline, early organic dementia, severe mood disturbance, or other conditions associated with neuropsychological impairment" (Hanna et al., p. 176). The authors believed that neuropsychological impairment contributed to poor clinical performance. Hanna and colleagues suggested that identifying irreversible neuropsychological impairment before the extensive three year training program could have saved the physicians embarrassment and the expense of the training. Also, reversible neuropsychological impairment could have been addressed early, increasing the likelihood of recovery.

As a result of the study by Hanna et al. (2000), the PREP program began to include neuropsychological assessment as a part of the intensive one-day evaluation. In another study conducted at PREP, Turnbull et al. (2000) reported that of 27 physicians assessed during a one year period, 7 demonstrated moderate to severe cognitive problems. Therefore, a significant minority of physicians evaluated (26%) demonstrated a problematic level of neuropsychological impairment. This finding suggests that neuropsychological assessment "may address areas relevant to the practice of medicine" (Turnbull et al., p. 179). One problem with this study is the lack of a control group; therefore

it has not been demonstrated that there would be more or less cognitive impairment in a group of physicians not referred for competency evaluations.

There are only a few studies that have explored the relationship between neuropsychological assessment and physician clinical performance. However, the early research seems to suggest that there is a significant relationship between neuropsychological measures and clinical performance. For example, Schueneman, Pickleman, Hesslein, and Freeark (1984) looked at the relationship of neuropsychological measures of cognition and perceptual-motor abilities to surgical skill. The authors found that visual-spatial problem solving and the ability to identify essential from nonessential details predicted surgical skill.

Another study that also demonstrated a significant relationship between neuropsychological measures and clinical performance was conducted by Reich et al. (1999). In this study, tests of vigilance, attention, and speed of processing were significantly related to the clinical performance of anesthesiology residents. Both of these studies support the intuitive conclusion that neuropsychological test performance is associated with physicians' clinical performance.

The results of the above studies demonstrate that neuropsychological assessment is an important part of understanding physician competency. Another study adds weight to this conclusion. Williams, Williams, and

Norcross (2002) conducted a study comparing the neuropsychological performance of physicians disciplined by a medical board and involved in competency evaluations, with physicians whose competency was not being questioned. While the sample size was very small (5-disciplined physicians and 9 non-referred physicians), results of the study indicated that physicians who were involved in competency evaluations performed significantly worse on three tests of cognitive ability and achievement than physicians who were not involved in competency evaluations. This suggests that there may be a difference in the cognitive picture of physicians who are referred for competency evaluations and a control group of physicians.

The PACE study (Williams et al., 2002) was the only study known to this author that included some type of physician control group. However, their sample size was very small (N = 14). Thus, the lack of control groups or sufficiently large control groups in the research leaves a gap in the empirical support for there being differences in the results of cognitive assessment between physicians who have been referred for competency evaluations and physicians who have not been referred for competency evaluations. This current study incorporated a larger control group to determine if, indeed, there were cognitive differences between these two physician groups.

Data for the group of physicians who have been referred for competency evaluations was obtained from The Center for Personalized Education for

Physicians (CPEP) program. This program provides competency evaluations for physicians from all over the United States. Physicians are referred through hospital peer review committees, state medical boards, medical groups, physician health programs, health attorneys, and disability insurers. A small portion of physicians evaluated are self-referred (Stulp deGroot, Barely, Dickerman, & Korinek, 2004). CPEP has been evaluating and providing education for physicians since 1990. Beginning in 1997, a computerized neuropsychological screening instrument, The MicroCog: Assessment of Cognitive Functioning (Powell et al., 1993), became a part of the standard battery of tests (E. Korinek, personal communication, February 22, 2005).

The MicroCog: Assessment of Cognitive Functioning (Powell, et al., 1993), originally called the Assessment of Cognitive Skills (ACS), was developed as a neuropsychological screen to detect mild cognitive impairment in physicians in hopes of reducing malpractice liability (Elwood, 2001). As a result of this underlying goal, the MicroCog (Powell et al.) was designed to have practically no ceiling effect, making it ideal to identify cognitive deficiencies in highly educated and functioning individuals. The instrument was designed to simplify the human-computer interaction, thus reducing the impact of previous computer experience on the outcome (Powell et al.). The psychometric properties of the MicroCog (reliability and validity) fall into the acceptable to excellent range. "Overall, MicroCog was normed better than most

neuropsychological tests when it was introduced in 1993 and, despite minor criticisms, still compares favorably with current neuropsychological test batteries " (Elwood, 2001, p. 92).

One problem with this assessment instrument is that its current version was not normed on a physician population, as was the original version, the ACS. When the Psychological Corporation purchased the instrument, its name was changed, along with some test items (Douglas Powell, personal communication, November 4, 2003), and it was normed as a general neuropsychological screen (Kane, 1998). Therefore, the current norms include only three education categories less than a high school education, a high school education, and more than a high school education (Powell et al., 1993). The mean educational level for the group labeled greater than a high school education is 15.6 years. Since most physicians have a minimum of 22 years of education, the current educational levels may not provide a valid comparison group for physicians. This study collected data on the MicroCog (Powell et al.) from a physician population that has not been referred for competency evaluations in order to provide a normative group that could be compared to a group of physicians who have been referred for competency evaluations.

The fact that the MicroCog (Powell et al., 1993) is a computer-based assessment raises the question of how suitable computerized assessments are as neuropsychological screens. Computerized assessment has evolved to the point

where appropriately developed programs have the sensitivity, specificity, and psychometric properties to assess for neuropsychological impairment (Fray, Robbins, & Sahakin, 1996; Green, Green, Harrison, & Kutner, 1994). In some cases computer-based neuropsychological instruments are more sensitive to cognitive deficits than traditional neuropsychological measures (Fowler, Saling, Conway, Semple, & Louis, 1997).

Statement of the Problem

The purpose of this study was to provide more information about the cognitive differences between physicians referred for competency evaluation and a control group of physicians, who had not been referred for a competency evaluation. The incorporation of a physician control group provided a normative group that could be compared to the competency evaluation physician group to determine if they demonstrated more cognitive difficulties than the physician control group.

Another purpose of this study was to determine if a control group of physicians performed higher than the age- and education-corrected norms currently available on the MicroCog (Powell et al., 1993).

Hypotheses

Hypothesis #1. As measured by the MicroCog (Powell et al., 1993), the physician group referred for competency evaluations will score significantly lower than the control physician group (those who have not been referred for

competency evaluations), on measures of speed, accuracy, and cognitive proficiency.

Hypothesis #2. The proportion of physicians with cognitive impairment in the group of physicians referred for competency evaluations will be significantly greater than the proportion of physicians with cognitive impairment in the control group of physicians (physicians who have never been referred for a competency evaluation).

Hypothesis #3. As measured by the MicroCog (Powell et al., 1993), the magnitude of difference on mean scores of the General Cognitive Proficiency Index between the two physician groups will increase as physician age increases.

Hypothesis #4. The control physician group will score significantly higher than the age-corrected norms at the highest educational level for the MicroCog (provided in the manual; Powell et al., 1993) on measures of speed, accuracy, and cognitive proficiency.

Study Definitions

In discussing physician impairment, it is imperative to have a working knowledge of a few terms that were used in this current study.

Impaired physicians - According to the American Medical Association ("The Sick Physician," 1973) an incompetent or impaired physician is "one who is unable to practice medicine with reasonable skill and safety to patients

because of physical or mental illness, including deterioration through the aging process, or loss of motor skill or excessive use of drugs including alcohol" (p. 684). A physician could be impaired for a variety of reasons; some possibly related to cognitive problems and some with no relationship to cognitive problems. Therefore, this term refers to any physician who places patients at risk.

Cognitively impaired persons - Individuals or groups of individuals who have demonstrated a level of cognitive deficits such that they are identified by neuropsychological test criteria or neuropsychologists as functioning in a cognitively deficient range. Please note that several studies mentioned in this literature review did not provide a definition for cognitive or neuropsychological impairment.

Impaired performance on the MicroCog (Powell et al., 1993) - An impaired performance on the MicroCog was determined by a General Cognitive Proficiency Index score of greater than one and one-half standard deviations below the mean (standard cutoff score = 77.5), or any two index scores that were one-half standard deviations below the mean. It is important to note that an impaired performance on a neuropsychological screen should not be used as a definitive clinical evaluation. This instrument is a screen and impaired performances should be followed up with a full neuropsychological evaluation to determine if an individual has cognitive impairment.

Summary

This chapter provided a brief overview of how neuropsychological measures help in the evaluation of impaired physicians. First, factors that contribute to and mitigate physician impairment were identified. Next, a discussion of how medical boards and physician evaluation programs address the problem of physician competency was presented, followed by information on how neuropsychological assessment has come to be an integral part of these evaluations. A brief review of literature that suggests a relationship between neuropsychological measures and physician clinical performance was presented. This was followed by a summary of the MicroCog (Powell et al., 1993), a neuropsychological screening measure which is currently being used by a physician evaluation program. Finally, the purpose of this study was discussed, the hypotheses were presented, and definitions were provided.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

This section is comprised of an overview of physician impairment, a presentation of cognitive difficulties that could contribute to physician impairment, a discussion of the importance of neuropsychological assessment in physician competency evaluation, a brief survey of the history and development of computer-based neuropsychological assessments, and a review of the MicroCog: Assessment of Cognitive Functioning (Powell et al., 1993).

Physician Impairment

Physicians hold a unique position of trust in our society. Individuals trust physicians with not only intimate aspects of themselves, but ultimately with their lives (Gendel, 2000). Because of this trust, "the public expects the medical profession to be accountable and, in particular, to monitor and deal with the competence of individual practitioners" (Trunkey & Botney, 2001, p. 393). According to the American Medical Association ("The Sick Physician," 1973) an incompetent or impaired physician is "one who is unable to practice medicine with reasonable skill and safety to patients because of physical or mental illness, including deterioration through the aging process, or loss of motor skill or excessive use of drugs including alcohol" (p. 684). This

definition of an impaired physician suggests that several factors could impact a physician's performance. Some behaviors associated with impaired physicians include unethical behavior, poor boundaries with patients, prescribing problems, documentation problems (Stulp deGroot, et al., 2004), poor patient outcomes, or not staying current on research, treatment, and medication.

According to the American Medical Association (Pasko & Seidman, 2002) there were 813,770 practicing physicians in the United States in 2000. Some estimates suggest that at any one time 7% to 10% of all physicians are impaired enough that their practice of medicine is affected (Van Komen, 2000). Corroborating this estimate, Hall et al. (1999) found, through routine evaluations that used questionnaires filled out by several individuals associated with the physician, that about 9% of the physicians evaluated needed further assessment. While 7% to 10% may appear to be only a small proportion of the total physicians practicing, if these estimates are accurate, it would mean that the actual number of impaired physicians in the United States in 2000 was between 57,000 to 80,000.

This number justifies concern regarding competence that physicians themselves are expressing (Epstein & Hundert, 2002; Trunkey & Botney, 2000).

Epstein and Hundert (2002), who are both physicians, expounded on the definition of what it means to be a competent physician. They proposed "that professional competence is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection

in daily practice for the benefit of the individual and community being served" (p. 266). Through this insightful definition, it is clear that the occupation of being a physician demands high and complex levels of cognitive functioning. Minor changes in cognitive functioning that might otherwise go unnoticed in most individuals, could significantly impact a physician's ability to provide competent care (Thompson, 2003). Therefore, professional competence is a dynamic condition that requires constant vigilance on the part of physicians as they acquire competence in medical school and residency, and work to maintain it throughout the time of active practice (Epstein & Hundert).

Cognitive Factors that Could Contribute to Physician Impairment

Poor cognitive functioning is another factor that could contribute to impaired physician performance. However, cognitive functioning as a contributing factor to physician impairment frequently has been ignored and supported by very little research. Thus, this current study focused solely on the contribution that cognitive impairment might have on physician impairment. This section will review some of the most common threats to physician competency that are related to cognitive functioning, such as sustained alcohol and drug abuse, neurological insults, normal aging, disease, and sleep deprivation.

Alcohol and Drugs

Approximately 75% of the physicians who appear before state medical boards are there due to problems related to drug or alcohol abuse (Clay &

Conaster, 2003; Van Komen, 2000). Alcohol is the most often used substance by physicians (Hughes et al., 1992). Estimates for alcohol abuse range from 7% to 13% of practicing physicians (McAuliffe et al., 1991; Moore, Mead, & Pearson, 1990; Regier et al., 1990). Based on the results of an anonymous questionnaire survey (with fixed alternative items) mailed to physicians, pharmacists, and medical students randomly selected in the state of Massachusetts, McAuliffe et al. found that about 10% of the physicians who responded self-reported as heavy drinkers, defined as drinking 61 to 90 drinks per month. Almost 4% of the responding physicians admitted to drinking 91 or more drinks per month. McAuliffe et al. also noted that since their data were based on self-report, the actual prevalence of alcohol abuse is probably higher, due to the potential for denial of problems with alcohol. Alcohol abuse does adversely impact a physician's ability to practice as noted by the number of charges brought to state medical boards related to alcohol use (Clay & Conaster; Van Komen).

Certain types of drug abuse occur more frequently in physicians than in the general population. This is partially due to the availability of drugs to physicians. While physicians as a group tend to be less likely to abuse illicit drugs, they are more likely to use prescription medication than the general population (Hughes et al., 1992). Approximately 4% of physicians in Hughes et al.'s study reported having a drug abuse or dependence problem in their

lifetime. The two most commonly abused medications in that study were opiates and benzodiazepine tranquilizers.

Many physicians who abuse substances eventually find that their ability to practice medicine safely is impaired. There is the concern of practicing while intoxicated, along with the concern that long term substance abuse could adversely impact cognitive functioning even when a physician is sober. Physicians who have substance abuse problems threaten the safety of their patients. However, substance abuse is not the only cause of physician cognitive impairment.

Normal Aging and Diseases that Impact Cognitive Functioning

"Physicians, like anyone else, are subject to the same kinds of cognitive deficits resulting from tumors of the brain, cerebral vascular disease, head injury, Alzheimer's disease, demyelinating disease, Parkinson's disease, brain infections, and toxic and metabolic brain disorders, ... and aging" (Madden, 1988, p. 201). There is virtually no information on how many physicians have diseases that adversely impact their cognitive functioning. The closest estimates include all physical disabilities, from having a leg amputated to diseases that directly impact cognitive functioning. The prevalence of physical disabilities in the physician population is estimated to be between 2.5% and 4% (Lewis, 1983; Martini, 1987). Madden (1988) reported that 17.5% (4) of the 70 physicians seen by the Physician Rehabilitation Program for the State of Maryland in 1986 had physical disabilities (White & Hayes, 1987). However,

the prevalence of disabilities that adversely impact cognitive functioning is unknown.

"Aging may create changes which make it increasingly hard for the physician to maintain the standard of practice" (Sadavoy, 1994, p. 265). There are several aspects of reduced cognitive functioning that are associated with normal aging. Approximately 17.8% or 144,939 practicing physicians were 65 and older in 2000 (Pasko & Seidman, 2002). In 1970 only 11.5% of physicians were 65 and older, suggesting that the number and proportion of practicing older physicians is increasing. Physicians tend to live longer and work longer than the overall population, leaving them vulnerable to the effects of aging on their practice (Lunsford, 1981; Sadavoy). The rest of this section will explore the cognitive impact of normal aging.

There are several mitigating factors that need to be considered in developing an understanding of normal aging and related cognitive decline. One of these factors is that as humans age, individual differences in cognitive abilities increase (Christensen, 2001; Nelson & Dannefer, 1992; Schaie & Willis, 1993; Wilson et al., 2002). Nelson and Dannefer examined 185 studies with age-based generalizations and found that about two-thirds of the studies reported increases in variability with increasing age. Therefore, the variability in a group of 60 year olds is greater than for a group of 30 year olds. This greater variability is found in cognition, physiology, and personality. Increasing variability with age is also true for the physician population (Powell

& Whitla, 1994). This means that it is less helpful to use measures of central tendency to describe the cognitive functioning of older individuals. It is probably more helpful to use individual performance scores to characterize the functioning of older individuals. Thus, it is important for the reader to remember "that change in cognitive function in old age is highly specific to the individual" (Wilson et al., p. 190).

Researchers have found several factors that impact variability in cognitive abilities with increasing age. First, having a greater number of years of education seems to reduce the impact of age-related cognitive decline (Christensen, 2001; Compton, Bachman, Brand, & Avet, 2000; Maddox, 1987). The 22+ years of education that are required of physicians should reduce the risk of functional impairment initially and over time. "Adequacy of economic resources" is also associated with reduction in the risk of functional impairment (Maddox). This variable is also significant for physicians, since most physicians have adequate financial resources. Finally, two factors that contribute to successful cognitive aging are increased activity and good health, both of which physicians as a whole experience (Christensen). Thus, as a group, physicians have the benefit of factors that contribute to reduction in the risk of functional impairment.

While these "protective" factors may slow or alter the pattern for age-related cognitive decline, older physicians still experience cognitive decline as they age. With age comes a decline in motor response time, ability to sustain

concentration over long periods of time, the ability to sort out essential versus non-essential details, visual-spatial abilities, and the ability to multi-task. As humans age, it becomes more difficult to learn new material in a short amount of time (Goldstein, 2000; Powell & Whitla, 1994). There is also decline in some memory functions. Wilson et al. (2002) found age-related declines in episodic memory. Another study found that decline in memory is not uniform (Zelinski & Kennison, 2001). In this study the greatest decline was in text recall followed by item recall. The authors concluded that any sharp decline in memory functioning before the age of 80 should be viewed as abnormal. While these studies did not look at physicians specifically, it is intuitively clear that cognitive decline in many of these areas could adversely impact a physician's ability to provide safe and effective medical care.

There are studies that have researched the impact age has on different aspects of physician performance. In studying physicians who had been referred for competency evaluations, age was found to be the variable that was most predictive of performance. Older physicians performed less well than younger physicians on cognitive exams, review of patient charts, following acceptable standards of patient care, oral exams, meeting with "standardized" patients, and possibly patient outcomes (Choudhry, Fletcher, & Soumerai, 2005; Norman et al., 1993). Several other studies support this finding. Norcini, Lipner, Benson, and Webster (1985) found that older internists performed worse than younger internists on a re-certification examination. McAuley, Paul,

Morrison, Beckett, and Goldsmith (1990) studied the results of a five year peer review program where physicians were randomly assessed by their peers as part of a campaign to demonstrate accountability in Canada. The results showed that increased age was significantly related with lower standards in record keeping, negatively impacting patient care. Of the physicians who were evaluated over the age of 75, 35% demonstrated vastly insufficient records or unacceptable levels of patient care. Sixteen percent of physicians evaluated in the age group of 50 to 74 had unacceptable patient records and poor patient care, while 9% of physicians under the age of 50 demonstrated these same deficits. Choudhry et al. reviewed 62 studies that contained results about quality of care and physician age. These authors concluded that "Physicians who have been in practice longer may be at risk for providing lower-quality care" (p. 260). From these studies it is clear that some physicians decline in performance and in ability to safely practice medicine as they age.

Until recently, cognitive or neuropsychological difficulties have not held a prominent position in the on-going discussion about physician impairment. With continued research on the neuropsychological characteristics of physicians, it is hoped that cognitive functioning will become an essential component of any competency evaluation.

Regardless of the etiology, physician impairment is a substantial problem for public safety. As noted above, up to 80,000 physicians who were practicing in 2000 were impaired to the point that they were unable to provide

appropriate medical care to patients (Hall et al., 1999; Pasko & Seidman, 2002). The medical profession has long struggled with how to identify and treat this impaired group. State medical boards are supposed to identify physicians at risk and intervene before the public is at risk. The weakness in this system is in identifying the physicians at risk (Clay & Conaster, 2003). State boards only discipline about .24% (Morrison & Wickersham, 1998) to .37% (Clay & Conaster) of the total number of physicians practicing. Very few referrals, about 3% of total referrals, originate from within the physician community (Clay & Conaster). One of the reasons for this minimal reporting by physicians is that they have to worry about being sued for libel or slander if they report another physician (Chapman, 1983).

Another reason physicians do not report incompetent physicians is related to the reason physicians do not recognize their own weaknesses. The culture of being a physician is fraught with the need to appear healthy and competent (Sargent, 1985). Therefore, from medical school onward, physicians learn to deny and ignore their own and colleagues' weaknesses, errors, and incompetence. The end result of this physician culture is that it is very unlikely that physicians will independently seek out treatment for themselves (Sargent), or report other physicians (Terry, 2002).

Physicians and Neuropsychology

In her chapter on neuropsychological assessment of physicians, Thompson (2003) cited Kapur (1997) in discussing what needs to be done when

physicians are ill or have injuries that affect their cognitive functioning and ability to practice medicine safely. Kapur called for the early involvement of a neuropsychologist when there is suspected or clear decline in cognitive functioning. A neuropsychological consultation could provide the physician and the people working with him or her a clearer understanding of the cognitive symptoms that may be impacting the physician's ability to function in the work environment. Early neuropsychological evaluation could lead to counseling or therapy to address some of the cognitive symptoms and to facilitate coping with these symptoms in the work environment. Kapur concluded with a powerful reason why early neuropsychological assessment benefits the physician. He said, "These types of measures will help to ensure that the self-respect and mental well-being of the physician are maintained, and that matters such as career development, financial affairs, etc. can be discussed on a rational basis" (p. 402). Early intervention also protects the welfare of patients being seen by potentially impaired physicians, as well as the physicians themselves.

In considering the problem of early identification of cognitively impaired physicians, some physicians have recommended that regular evaluations, including neuropsychological assessment, be conducted on physicians after they reach a certain age (Trunkey & Botney, 2000). These authors compared the airline industry standards for pilots to the medical industry standards for physicians. Pilots are required to pass a first class medical examination every six months. These medical evaluations check for

any possible neurological insults, diseases, or impairment. Pilots are also evaluated on a flight simulator every year. These tests are supplemented with random assessments throughout the year. Pilots employed by commercial passenger airlines are required to retire at the age of 60, with no exceptions. Trunkey and Botney suggested that physicians receive a modified version of the airline pilots' regimen for accountability to public safety. At 50 years of age, physicians would be required to obtain a medical certification exam every two years. The authors recommended that the exam include a neuropsychological screen and suggested the MicroCog as a possible instrument (Powell et al., 1993). At the age of 60, physicians would be required to complete a mandatory evaluation every year. This yearly evaluation would allow highly functioning older physicians to continue practicing with confidence. The yearly evaluation would also catch early cognitive decline and provide the opportunity for treatment and possible reversal of cognitive deficits (Trunkey & Botney). Finally, this type of evaluation could provide physicians with information about when it would be appropriate to retire from practice. The next section reviews how neuropsychological test performance relates to physician clinical performance.

Research on Neuropsychological Measures and Physician Performance

There is very little research on the ecological validity of neuropsychological tests for physician clinical performance (Thompson, 2003). However, the few studies that do address this issue suggest that

neuropsychological test performance is related to physician performance. This section includes a discussion of the research that deals with the relationship between neuropsychological measures and physician performance. This segment also includes information from studies on the performance of sleep-deprived physicians, a condition that may create temporary cognitive decline. A final section considers the limitations of the research available, along with the limits of the instruments used to conduct valid research.

In the earliest study, Schueneman, Pickleman, Hesslein, and Freeark (1984) wanted to "identify neuropsychologic measures of cognitive and perceptual-motor abilities related to surgical skill..., to evaluate the relative importance of these abilities in accounting for variation among residents, and ... to compare the efficiency of neuropsychologic variables with academic indices in the prediction of surgical performance" (p. 289). One hundred twenty medical residents participated in the study. Each participant was given a battery of neuropsychological tests, with a particular focus on visual-spatial abilities, motor sequencing, and fine motor coordination. Residents' performance was evaluated by attending surgeons using 12 rating scales with content relevant to components of surgical techniques. The neuropsychological tests explained 46% of the variance in predicting surgical performance, while the scores on the Medical College Admission Test only accounted for 12.6% of the variance. The most significant neuropsychological predictors were visual-spatial problem-solving and the ability to determine essential from nonessential details.

In another study, Reich et al. (1999) looked at the relationship of cognitive, personality, and academic variables to clinical performance in anesthesiology residents. Sixty-seven participants completed a series of personality, academic, and cognitive assessments within 13 months. The neuropsychological assessment included a continuous performance task that required response to a visual stimulus while distracting sounds were played on a tape recorder. This was done to simulate the normal background noise of an operating room. The Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977) was used to assess divided and sustained attention and speed of information processing. Administration of the PASAT (Gronwall) involved digits being verbally presented and the participant being asked to add the last two digits given for a series of 50 digits. This study used three sets of 50 random digits, which were presented at an increasing rate of speed. To assess clinical performance, residents were rated by attending physicians on a monthly basis. A low score on the third and most difficult trial of the PASAT (Gronwall) and a high number of commission errors on the Vigil (Cegalis & Bowlin, 1995), a sustained attention visual target task, were predictive of poor clinical outcome (Reich et al.).

Studies of sleep deprivation and physician performance on cognitive clinical variables provide some insight as to how an over-taxed brain impacts performance. In their review of sleep deprivation studies, Weinger and Ancoli-Israel (2002) stated that the literature on the negative impact of sleep

deprivation for non-medical settings is conclusive, in that sleep-deprived individuals tend to perform 1.4 standard deviations below individuals who experienced normal sleep. However, the evidence for physicians in medical settings has not been as conclusive. The authors claim that the studies from 1985 to 1992 had significant methodological problems. Some of these methodological limitations included "inadequate controls, no randomization, inadequate consideration of training effects over time, or failure to control for circadian effects" (Weinger, & Ancoli-Israel, p. 956). More recent studies have begun to suggest that sleep deprivation does negatively impact physician performance (Weinger and Ancoli-Israel). The following section includes a review of some of these more recent studies.

Taffinder, McManus, Gul, Russell, and Darzi (1998) studied the effects of sleep deprivation on surgical skills. Residents were assessed under three conditions; an undisturbed night of sleep, a night of sleep with three interruptions (midnight, 3 AM, and 6 AM), and a night with no sleep. Surgical skills were assessed using a virtual-reality laparoscopic surgery simulator. Surgeons who were awake all night made more errors (20% more) and took longer (14% longer) to complete the tasks than surgeons who slept all night (Taffinder et al.). These results were corroborated in another study (Grantcharov, Bardram, Funch-Jensen, & Rosenberg, 2001) that assessed surgical dexterity after a night of less than three hours of sleep. In this study, surgeons in training performed a virtual reality laparoscopic cholecystectomy

after a night of interrupted sleep and a night of full sleep. During the condition of limited sleep, physicians took significantly longer and made more errors than physicians with a night of uninterrupted sleep. A weakness in both of these studies is that the physicians were not performing actual surgeries, which could have had an arousing affect on the operating physician.

In another study by Wesnes, Walker, and Walker (1997) physicians' cognitive abilities and emotional state were assessed after a weekend on call and after a weekend off call. Cognitive abilities were evaluated by a computerized cognitive assessment program designed to assess attention, working memory, and long term memory. After a weekend on call physicians' performance was adversely impacted. Attention speed, working and long-term memory, and vigilance scores were significantly lower after a weekend on call, when compared to a weekend off duty with full nights of sleep. Also after a weekend on call, physicians felt less energetic, more confused, and less confident. Robbins and Gottlieb (1990) also used cognitive assessment to evaluate resident physician performance after a night on call and a night of full sleep. After a night on call, obtaining an average of 2.1 hours of sleep, physicians demonstrated significant deterioration on measures of visual vigilance, hand-eye coordination, and the ability to process complex data. Supporting the results of this study, Murray and Dodds (2003) conducted a study to assess vigilance in trainees and consulting anesthetists after a night on call. A driving simulator was used to assess vigilance. Participants took the

driving test twice, once to provide a baseline score, and again after either a night of sleep or a night on call. There was a significant difference in the change of scores on the simulator between the group that had a night of uninterrupted sleep and the group that was on call. After a night on call, participants' scores on the vigilance task showed an increase in steering errors and the undisturbed sleep group showed a decrease in steering errors (Murray & Dodds).

The final study was conducted with experienced physicians and not residents (Smith-Coggins, Rosekind, Hurd, & Buccino, 1994). Attending physicians' cognitive and motor abilities were assessed after a night of sleep and daytime work, and then again after daytime sleep and nighttime work. The results of this study suggest that physicians slept less during the daytime sleep compared to the nighttime sleep. Also, physicians working on the night shift were slower at intubating a mannequin and committed more errors as the night progressed. While the question of how sleep deprivation impacts physician performance is still undecided, the above studies do suggest that cognitive and clinical performance is adversely impacted when participants are sleep deprived. The sleep deprivation studies provide examples of how over-taxed brains could impact physician performance, both cognitively and clinically.

Although there is a paucity of information on the relationship between physician clinical performance and cognitive measures, early studies suggest that there is a significant relationship. More research is needed to determine

what cognitive measures predict physician performance. Also lacking is the availability of physician-based norms on neuropsychological tests (Thompson, 2003). However, "Given our current state of knowledge, neuropsychological evaluation appears to be useful in the evaluation of physicians whose competency to practice is being questioned" (Thompson, p. 378).

Thompson (2003) goes on to list what neuropsychological assessment can contribute to the evaluation of physicians:

- Detection of subtle (but professionally significant) impairments in highly functioning professionals that are not picked up by other evaluation methods
- Evaluation of behavioral manifestations of neurological disease (e.g., Parkinson's disease), systemic illness (e.g., chronic obstructive pulmonary disease), or substance abuse/dependence for the purpose of determining safety to practice medicine
- Monitoring of mental functioning in the presence of expected recovery from illness (e.g., stroke) or injury (traumatic brain injury) to help determine whether and when return to practice is safe and appropriate
- Monitoring of progression of cognitive impairment in physicians with chronic illnesses (e.g., multiple sclerosis) to determine whether and when the physician becomes unsafe to practice medicine
- Differentiating depression from dementia in older physicians (p. 379)

This list and previously cited research suggest that neuropsychological assessments have much to offer physicians and physician programs that have the responsibility of supervising the safe practice of physicians.

Programs for Evaluating Physician Competency

According to Thompson (2003) there are three paths that lead to physicians being evaluated by a neuropsychologist. These include referral by a treating physician, referral through a physician health program, or referral

through a physician competency evaluation. This current study was concerned with the physician competency evaluation. There are physician assessment programs that provide evaluations of physician competency throughout North America. This section reviews two of these programs and the research that is being conducted through these programs. Also, the program with which this study is associated is described.

The first program is the Physician Review Program (PREP) based in Canada. This program was developed through the College of Physicians and Surgeons of Ontario (CPSO), the licensing authority for Ontario. This group has implemented a randomized peer review program for the 26,000 physicians in the province (Hanna, Premi, & Turnbull, 2000). Approximately 8% of the physicians reviewed demonstrated serious deficiencies (McAuley & Henderson, 1984). These physicians were given the opportunity to correct these deficiencies and were then re-evaluated with a follow-up review. If a physician failed to demonstrate improvement during the follow-up evaluation, or there was a concern about the competence of the physician, a referral could be made to the PREP evaluation program (Hanna, Premi, & Turnbull). The evaluation program provides an intensive one-day evaluation of physicians. The evaluation includes multiple choice questions, peer-observed encounters with patients, chart-simulated recall, and structured oral examinations. The results of the examinations are then integrated into a global composite score and physicians are placed into one of six categories of competence. Categories I

and II indicate that the physician demonstrated few or no deficiencies. Categories III and IV are comprised of physicians who demonstrated moderate to major problems and whose competence is seriously questioned. Categories V and VI are comprised of physicians who are either unsafe to practice without supervision or who are unsafe to practice under any condition. Physicians evaluated by the PREP program, and for whom it was determined that remediation was needed, were referred to physician remedial continuing medical education (CME) as a part of their individualized educational plan. Many physicians' performance improved after participating in a self-directed CME (Hanna, Premi, & Turnbull).

Two studies that demonstrate the need for neuropsychological assessment as a part of these intensive assessments have come out of the PREP program. The first study (Hanna, Premi, & Turnbull, 2000) involved the selection of five moderately to seriously incompetent physicians, having been identified as PREP level III or IV. It is important to note that a neuropsychological evaluation was not a part of the PREP evaluation. These physicians' performance did not improve after participating in the self-directed CME. Therefore, the CPSO developed a concentrated and extended educational program for these five physicians. The extended program lasted for three years and incorporated individualized and small group training sessions. The results of this intensive program were that one physician obtained an improved PREP level, one physician retained the same PREP level, two physicians' PREP level

declined by one, and one physician's performance declined by two PREP levels. In discussing the reasons for the lack of improvement among the majority of physicians in this program, the authors suggested that "it is possible that their incompetence arose from early age-related cognitive decline, early organic dementia, severe mood disturbance, or other conditions associated with neuropsychological impairment" (Hanna, Premi, & Turnbull, p.176). The authors added that, "If it could be demonstrated that at least some physician incompetence arises from irreversible cognitive difficulty, and if cognitive difficulty precludes remediation, then such physicians could be spared the embarrassment and expense of extensive and futile remediation, and in certain cases, they could receive treatment or appropriate disability coverage" (Hanna, Premi, & Turnbull, p.176). These authors made a strong argument for the inclusion of neuropsychological testing in physician competency evaluations.

In the second study from PREP (Turnbull et al., 2000), a neuropsychological evaluation was added to the intensive one-day assessment battery. The purpose of the study was to determine if cognitive difficulties contributed to the lack of success of physicians in the CME remedial program. The neuropsychological evaluation assessed problem solving, concept formation, reasoning, learning, memory, attention, complex mental tracking, verbal fluency, and mood. See Table 1 for a list of the neuropsychological instruments used in the battery.

Twenty-seven physicians were tested in a one year period. Of these 27 physicians, seven (26%) demonstrated moderate to severe cognitive problems; mood disturbance was a problem for an additional three (11%) physicians. Of the 19 physicians who had a PREP level of III or greater (moderate to major problems and competency is seriously questioned), six (32%) demonstrated moderate to severe cognitive difficulties. Therefore, approximately one-third of physicians who scored poorly on the PREP evaluation had cognitive difficulty. This suggests that cognitive testing "may address areas relevant to the practice of medicine" (Turnbull et al, 2000, p.179). One problem with this study was the lack of a control group with which to compare physicians being evaluated due to competency concerns. Thus, Turnbull and colleagues did not demonstrate that there was a difference between a "normal" group of physicians and the competency evaluation physicians.

The authors also suggested that neuropsychological testing is beneficial for identifying treatable conditions that adversely impact neuropsychological functioning. Early identification and treatment of reversible neuropsychological problems, such as cognitive problems caused by medication, nutritional deficiencies, psychiatric disorders, benign brain tumors, and chemical intoxication (Nixon, 1996), could allow physicians to continue practicing with confidence. Identifying physicians whose neuropsychological impairment is not treatable could prevent physicians from having to fail at attempts at educational remediation and also prevent costly lawsuits resulting

from impaired practice. These physicians could be treated with respect and understanding, and appropriate interventions such as help closing a practice and obtaining disability benefits could be initiated (Turnbull, et al.).

Turnbull et al. (2000) addressed the issue of using age-appropriate norms with physicians. Studies have demonstrated that both age and education impact an individual's performance (Heaton, Grant, & Matthews, 1991; Heaton, Chelune, Talley, Kay, & Curtiss, 1993; Powell et al., 1993). In normal adults, as age increases, cognitive performance decreases. Likewise, as physicians age, their performance typically will decrease. Therefore, adjustments to performance scores are made to correct for the decline in performance with increasing age. However, Turnbull et al. questioned this practice for physicians stating, "While adjusting these scores for age may be appropriate when clinically assessing cognitive deficits, such adjustments may be less helpful in assuring the quality of physicians" (p. 180). Practicing physicians need to maintain a certain level of cognitive functioning to continue to practice competently and safely (Turnbull et al.). Therefore, assessing physicians based on age-related norms may not provide a standard rigorous enough to determine a physicians' ability to practice competently.

The second program that conducts physician evaluations was developed by the Department of Family and Preventive Medicine at the University of California in San Diego (Williams, Williams, & Norcross, 2002). The Physician Assessment and Clinical Education (PACE) program conducts a

Table 1 Neuropsychological Test Battery for PREP

Wechsler Adult Intelligence Scale-Revised Subtests (Wechsler, 1981; Similarities, Comprehension, Picture Completion, Picture Arrangement)
Wisconsin Card Sort (Heaton, 1993)
Rey Osterrieth Complex Figure (Rey-Osterrieth CFT; Rey, 1941, 1959)
Weschler Memory Scale-Revised Subtest (Wechsler, 1987; passages with 30-minute delay)
California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987)
Trail Making Test (Reitan, 1958)
Stroop Color Word Interference (Stroop, 1935)
Paced Auditory Serial Addition Task (Gronwall, 1977)
Verbal Fluency (Benton & Hamsher, 1978)
Profile of Mood States (McNair, Lorr, & Droppleman, 1981)

Note. From "Cognitive Difficulty in Physicians," by Turnbull et al., 2000, Academic Medicine, 75(2), p. 178. Copyright 2000 by Lippincott, Williams, & Williams. Adapted with permission.

two-day intensive assessment of physicians who have been disciplined by the California State Board of Medical Examiners. The first day of assessment is comprised of a physical exam, the physician's assessment of a mock patient, and a clinical competency examination related to the physician's specialty area.

Table 2 Neuropsychological Test Battery for PACE

Wechsler Adult Intelligence Scale-Revised Subtests (Wechsler, 1981; Information, Digit Span, Vocabulary, Similarities, Picture Arrangement, Block Design)
Trail Making Test (Reitan & Wolfson, 1985)
Halstead Category Test (Halstead & Wolfson, 1993)
Symbol Digit Modalities Test (Smith, 1973)
Numerical Attention Test (Heaton, Thompson, Nelson, Filley, & Franklin, 1990)
Benton Multilingual Aphasia Exam (Benton & Hamsher, 1978; selected subtests)
Story Memory Test (Heaton, Grant, & Matthews, 1991)
Modified Complex Figure Test
California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987)
Wide Range Achievement Test-3rd Edition (Wilkinson, 1993)
Rorschach (Rorschach, 1921; used as cognitive assessment instrument)
MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989)

Note. Copyright 2003 from Neuropsychological assessment of physicians whose competency to practice medicine is being questioned by Thompson, L. L. In G. P. Prigatano, & M. H. Pliskin (Eds.), Clinical Neuropsychology and Cost Outcome Research: A Beginning. Reproduced by permission of Routledge/Taylor & Francis Books, Inc.

The second day includes a reasonably comprehensive neuropsychological evaluation (Thompson, 2003) that assesses the learning of verbal and non-verbal information, memory, vigilance, sustained attention, eye-hand coordination, ability to follow complex procedures, abstract and complex reasoning, and problem-solving abilities. See Table 2 for a list of the neuropsychological measures used during the second day of assessment. Based on the results of the assessment, an individualized clinical education program is developed for each physician.

The PACE program conducted a study designed to investigate the neurocognitive difference between physicians disciplined by a state medical board and involved in a competency evaluation, and non-disciplined physicians (Williams, Williams, & Norcross, 2002). Five physicians from the PACE program were compared to nine faculty members of the Department of Family and Preventive medicine. Both groups were given two subtests from the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1987; vocabulary and picture arrangement) and the Wide Range Achievement Test-Third Edition (Wilkinson, 1993). The physicians from the PACE program performed significantly lower on all three tests. The results suggest that neuropsychological deficits may occur more frequently in the population of physicians disciplined by governing boards. Since this is the only study with some type of physician control group, and the size of that control group was small, more research is needed to provide support for there being a difference

between physicians involved in competency evaluations and physicians not involved in competency evaluations.

More recently, the PACE program published another study that provided a description of the neuropsychological characteristics of 148 physicians that have been evaluated by their program (Perry & Crean, 2005). These authors excluded physicians whose primary language was not English, citing Jacobs et al. (1997) who reported the difficulties associated with neuropsychological assessment of individuals whose primary language was not English. The PACE physician group had a mean age of 54.4 years, with a standard deviation of 10.1. Males comprised 93.2% of the sample. For medical specialties, 37.2% were primary care; 25.7% were surgical specialties (includes OB/GYNs); and 37.1% were in the "other" category (includes anesthesiologists). Perry and Crean reported that 51% of the physicians scored one standard deviation below the mean on four or more of the neuropsychological instruments. They also mentioned that 45% of the physicians scored one standard deviation below the mean on three or more subtests from the WAIS-R. As mentioned by the authors, a weakness of this study was the lack of a control group of physicians with which to compare the competency evaluation physician group. Instead, they compared these physicians with reference normative samples. The authors reported that the competency evaluation physicians demonstrated deficits on tests of attention, complex reasoning, learning, eye-hand coordination, and sequential processing. Perry and Crean suggested that these physicians, "when

faced with complex and novel situations in which a quick, and yet complex, decision is warranted, they may be prone to poor problem solving and attention to critical detail, rendering them vulnerable to errors in judgement" (p. 169). These authors proposed that in order to make any conclusive statements about cognitive differences between competency evaluation physicians and "non-impaired" physicians, an age-matched physician control group would need to be included in the research. This current study included such a group.

Finding similar results, Madden (1988) reported results from a two year time period of evaluating physicians in Maryland. A total of 11 physicians were evaluated. In this program neuropsychological testing along with practice competency exams and personality tests were given to physicians. To assess neuropsychological functioning, the complete Halstead-Reitan Neuropsychological Test Battery (Reitan & Wolfson, 1985), the WAIS-R (Wechsler, 1981), and the Wide Range Achievement Test (WRAT; Wilkinson, 1993) were administered. Madden reported that 8 of the 11 physicians tested had Full Scale IQ scores in the average or borderline range. Seven (63.6%) physicians scored in the impaired range on the neuropsychological battery. Madden pointed out that some physicians scored within acceptable ranges on the IQ test and the WRAT, but were found to have significant neuropsychological deficits when administered the neuropsychological tests. As a result of this study, Madden expressed his emphatic support for neuropsychological testing for physicians being evaluated for competency. He

also recommended that the testing and reporting be done with "the greatest amount of gentleness and compassion" (p. 205).

The final program to be discussed is the Center for Personalized Education for Physicians (CPEP) program. This study incorporated data from this program. This program provides competency evaluations for physicians from all over the United States who have been referred through hospital review committees, state medical boards, medical groups, physician health programs, health attorneys, and disability insurers. A small portion of physicians evaluated are self-referred (Stulp deGroot, et al., 2004). Physicians referred to CPEP for competency evaluations are referred for a variety of reasons, such as concerns about clinical competence (35%), concerns about professional or interpersonal behavior (20%), documentation problems (13%), prescribing problems (9%), renewing, restoring and/or seeking licensure or privileges (8%), patient complaints (8%), illness or injury (5%), and other (3%), including self-referral (Stulp deGroot et al.). CPEP has been evaluating and providing education for physicians since 1990.

Originally, no neuropsychological screen was used in the competency evaluation; however, in 1997 a computerized neuropsychological screening instrument became a part of the standard battery of tests (E. Korinek, personal communication, February 22, 2005). The current CPEP evaluation consists of a two to three day intensive evaluation process. Information about the physician's performance is obtained through structured clinical interviews,

multiple choice questioning of knowledge, simulated patient exercises followed by chart writing, review of patient charts from the physician's practice, and computer case simulations. "This in-depth evaluation is tailored to the physician's specialty and practice, and provides detailed information about clinical competence in the areas of medical knowledge, clinical reasoning, documentation, communications, and cognitive function—while also identifying areas of educational need" (Center for Personalized Education for Physicians, n.d.). To determine if cognitive difficulties are contributing to the physician's competency problems, the MicroCog: Assessment of Cognitive Functioning (Powell et al., 1993), a computerized neuropsychological screen discussed in greater detail in a later section of this literature review, is administered to physicians being evaluated.

Thompson (2003) provided some descriptive statistics about the population of physicians assessed by CPEP between January 1997 and March 2000. Of the 110 participants included in the analyses, 25% (28) performed in a range that suggested significant cognitive impairment. These 28 physicians were encouraged to obtain a more comprehensive neuropsychological evaluation to better understand their cognitive deficits and strengths.

Thompson (2003) split the physicians who had received competency evaluations into two groups; the physicians who were encouraged to obtain further neuropsychological evaluation (referred group), and the physicians who did not appear to need further neuropsychological evaluation (not referred

group). Upon comparing the two groups, she found that the referred group was older, and that their mean performances on speed, accuracy, and proficiency were lower than the group that was not referred. For example, the mean percentile rank for accuracy was approximately the 19th percentile for the referred group, while the mean accuracy rank for the not referred group was approximately 49th percentile. This type of difference was also consistent for speed and proficiency. The difference between the two groups was even greater when medians were compared. Thompson also observed that the "better-performing" physicians, those not referred for neuropsychological evaluations, still performed below the 50th percentile on speed, accuracy, and proficiency when age corrected norms were used. She noted that the relatively low performance of the not referred physicians "may suggest that as a group, physicians referred to these evaluation programs are not functioning as well cognitively as the general group of practicing physicians" (Thompson, p. 387).

Thompson (2003) concluded that the field of neuropsychological evaluation of physicians is "ripe" for research. She stated that the lack of normative data for physicians in general is a basic concern. For some tests, normative data for individuals with high levels of education are available and currently used with physicians. However, on tests like the MicroCog (Powell et al., 1993), where the highest educational level is characterized as greater than high school education, the education levels are inadequate. Even with higher educational adjustment levels, research is still needed to determine whether

physician specific norms are required to serve as a comparison group. Thompson also suggested that research "to investigate the neuropsychological screening performance of physicians referred for physician-competency evaluations compared to a reference group of physicians who are practicing without difficulty" (p. 391) would be useful.

A Brief History of Computer-based Assessment in Neuropsychology

This section will provide a selective overview of the history of computers in neuropsychological assessment. The history begins with farsighted psychologists constructing awkward programs on equally awkward large mainframe computers. Both the computers and the programs were expensive and cumbersome. Eventually, the development of the microcomputer inspired a movement to computer-based neuropsychological assessments that were available to practicing clinicians. Based on the state of computerized assessments today, the future of neuropsychological assessment looks quite promising.

A significant marker in the history of the use of computers in neuropsychology occurred in the 1940s. Ward Halstead and his colleagues designed an automated console for testing patients whose occipital lobe had been removed (Halstead, Walker, & Bucy, 1940). While this "console" was not an actual computer, it demonstrated the idea that psychologists could develop automated systems to help assess individuals even before the digital computer was developed. However, "the invention in 1946 of the electronic digital

computer significantly and permanently changed the course of scientific inquiry" (Maulucci & Eckhouse, 1988, p. 557).

With the new digital computer that was capable of performing complex operations faster and more accurately than human beings (Maulucci & Eckhouse, 1988), psychologists began to change their understanding of how to assess individuals. In 1954, Meehl challenged psychologists to use actuarial methods to make predictions about human behavior. To facilitate the implementation of his actuarial methods, Meehl recommended the use of the Hollerith machine, a mechanical machine that tabulated data entered on punch cards. Despite the many psychologists who responded to his book with outrage, accusing Meehl of being "anti-clinical" (Meehl, 1996), actuarial methods and the computer were destined to become indispensable tools in the field of psychology and neuropsychology.

In the 1960s, neuropsychologists started to investigate the use of computers in assessment. For example, Knights and Watson (1968) described the development of a program that used scores from psychological tests to evaluate patterns of abilities and deficits in children who were cognitively impaired. Around this same time, Elwood (1969, 1972a, 1972b, 1972c, 1972d; Elwood & Griffin, 1972) developed an automated version of the WAIS. This automated WAIS was administered in a testing room designed specifically for this purpose. Once an individual was instructed by an examiner, he or she would press the "Ready" button and the test would begin. For several tests,

such as Block Design, drawers positioned around the testing console would pop out with the appropriate stimuli. The individual would then complete the task and return the drawer to its original location. Timing was based on how long the examinee had a drawer open. An examiner scored the test after the examinee left the room, by opening each drawer and scoring the results. Results were printed out on a teletype machine. Even though this version of the automated WAIS was a little crude, it demonstrated reasonable equivalence to the traditional paper and pencil version of the test (1969, 1972a, 1972b, 1972c, 1972d; Elwood & Griffin, 1972; Schatz & Browndyke, 2002). The 1960s were a time of creativity and foresight. However, the software and hardware available at that time created significant barriers in developing useful, reliable, and valid tests.

The 1970s were characterized by the development of computer programs that synthesized data from traditionally administered tests to determine if a person had brain damage or not, the location of the damage, and whether the damage was acute or not (Golden, 1987). One such program developed by Russell, Neuringer, and Goldstein (1970) was originally written in Fortran on punch cards and utilized a mainframe computer. These computer programs were reasonably successful in predicting the presence of brain damage. However, the computer programs were not as accurate as clinicians in making neuropsychological diagnoses (Heaton & Adams, 1987). Golden (1987) felt that the problem with these neuropsychological scoring programs

was that there were too many variables that contributed to the understanding of an individual's impairment that could not be synthesized by these programs.

The 1970s also saw the development of individual neuropsychological tests for the computer. Knights, Richardson, and McNarry (1973) compared the automated version of the Peabody Picture Vocabulary Test (Overton & Scott, 1972) and the Raven's Coloured Progressive Matrices Test (Raven, Court, & Raven, 1986) with the clinical administration of these tests. They found that children with low IQ's responded positively to the automated tests and reacted well to the touch screen system. There were no overall significant differences between the scores on the two versions of the tests. Another neuropsychological test adapted to the computer in the 1970s was the first version of the Category Test (Beaumont, 1975). This test included changing stimuli, response recording, and response latency timing (Schatz & Browndyke, 2002). Even though this version of the Category Test was not equivalent to the traditional Category Test, other versions that were equivalent to the traditional version soon followed (Choca, & Morris, 1992). While the 1970s were characterized by the development of many computerized neuropsychological tests, the requirements of the mainframe computers on which these tests were developed made these programs essentially inaccessible to most clinical practioners.

"The advent and proliferation in the 1980s of the desktop microcomputer made the advantages of the automated information processing

accessible to individual researchers" (Maulucci & Eckhouse, 1988) and clinicians. Therefore, during the 1980s many tests were adapted to or developed for the microcomputer. Several computer-administered neuropsychological assessment batteries were developed during this time (Branconnier, 1986; Swiercinsky as cited in Goldstein & Incagnoli, 1997). One of the first batteries was the SAINT by Swiercinsky (Golden, 1987). This battery included ten tests and was designed to be a comprehensive computer-administered neuropsychological battery. The test provided both auditory and visual stimuli, and allowed for various modes of input including a joystick, finger tapping, and multiple-choice items (Golden). While these test batteries were innovative and promising, Golden noted that there was a danger in assuming equivalency with the paper-and-pencil versions of these same tests. He believed that more research was needed to assess validity, reliability, and standardization norms.

Also in the 1980s, the evaluation of computer-based tests became more sophisticated. Researchers were no longer just giving two versions (computer and paper-and-pencil) of the same test to see if the results were similar. Pellegrino, Hunt, Abate, and Farr (1987) contributed to the development of methodological comparison in their assessment of 10 computerized tests of spatial ability. In this study, factor analysis played more of a role in determining exactly what computerized-based assessments were testing. These authors found that their computerized spatial tests generally demonstrated

excellent reliability, and (through factor analysis) that the computer spatial reasoning tasks measured the same abilities as the paper-and-pencil tasks.

The use of more sophisticated technology in the 1980s was evidenced by Crook and Larrabee (1988), who developed a creative computer-based assessment that replicated cognitive tasks of daily living (Maulucci & Eckhouse, 1988). The test was designed to simulate real-world memory tasks. For example, one task had the examinee view a video of a trip as viewed from the car. The computer then replayed the video, stopping at different points and asking the examinee which way to turn. Another task asked the examinee to memorize a telephone number. At a later time, the program asked the examinee to dial that number on a touch screen that resembles a telephone (Crook & Larrabee). With this innovative, reality-based program, Crook and Larrabee provided neuropsychologists with a glimpse of a very promising future for computer-based assessment. As Golden (1987) stated, "the use of computers in neuropsychology remains more of a promise than a reality....it is [also] clear that computers will play an increasingly large role in the field over the next decade" (p. 344).

The next decade, the 1990s, proved to be a time of continued growth in sophistication in computerized test development, which also led to increased test reliability and validity. One such development was in the area of timing procedures used by assessment programs. In their review of computerized neuropsychological test batteries, Kane and Kay (1997) discussed the need for

programmers to create and utilize more accurate timing mechanisms. There are different ways of programming how the computer will time an event. For example, the original version of the MicroCog (Powell et al., 1993) used a timing device that was inaccurate (Elwood, 2001). The creators of the MicroCog eventually replaced the old timing procedure with a new, more accurate program. It was hoped that this subtle change in the timing procedure could positively impact the reliability and validity of the MicroCog (Powell et al.). No research has been conducted to determine if reliability and validity were improved.

Other test batteries designed to detect cognitive impairment became available for the microcomputer in the 1990s. These batteries incorporated new software-based timing devices that were more accurate than the older timing procedures and a variety of input devices such as touch screens, light pens, and joy sticks (Kane & Kay, 1997). The psychometric properties of these new tests were also improved. Creators of these new test batteries have developed extensive norms, and worked on establishing validity and reliability. Factor analyses also provided information about what the batteries were assessing. After conducting a literature review of the various computer-based neuropsychological assessments available, Kane and Kay (1992) determined that these programs demonstrated adequate levels of sensitivity and specificity for use as screening instruments. Thus, the 1990s ended with computerized neuropsychological test batteries having achieved enough sophistication to be

used in clinical practice. The developments in the present decade promise to continue to improve the psychometric properties of these programs and to add new technology and creativity to assessment programs.

A new technology that has begun to appear in computer-based neuropsychological assessment and rehabilitation is virtual reality (Mehiltz, Kleinoedr, Weniger, & Rienhoff, 1998; Pugnetti et al., 1995). Elkind, Rubin, Rosenthal, Skoff, and Prather (2001) believe that virtual reality creates a safe environment that simulates real life settings with all of the distractions and interactions. Virtual reality provides a "life-like," three dimensional, highly interactive environment that adds ecological validity to a neuropsychological assessment.

Two other directions of development in computer-based neuropsychological assessment are the use of web-based screening tools and the assessment of concussions in athletes. Erlanger, Kaushik, Broshek, Freeman, Feldman, and Festa (2002) have developed a screening device that monitors cognitive status over time. This web-based program can be used at several sites simultaneously and can provide immediate feedback. Initial research into validity and reliability of this program has been positive. The second area of computerized test development is the assessment of sports-related concussions, which also incorporates the ability to monitor cognitive status over time. Schatz and Zillmer (2003) think that "computer-based assessment of sports-related concussion saves time, allows for team baseline testing, and can be

easily incorporated into the sports medicine environment" (p. 42). These authors view paper-and-pencil testing as too time consuming and cost prohibitive for the broad-based testing required by athletic programs. More and more professional and college sports teams are using computer-based neuropsychological assessments that are designed to evaluate the cognitive impact of concussions and to monitor cognitive recovery from sports-related head injuries.

As new computer technology is incorporated into innovative test design, test developers will be able to improve on accuracy, input and output modalities, speed, and human-computer interface. It is hoped that these improvements will lead to enhanced reliability and validity for computerized assessment. Thus, neuropsychologists will be able to take advantage of the benefits of computer-based assessment while minimizing its limitations.

Advantages of Computer-based Neuropsychological Assessment

With all of the advances in the technology of computers, the benefits of using computers in neuropsychological assessment continue to evolve. Some of the problems that were relevant in the 1960s, such as having to use mainframe computers, no longer exist. Computers have become faster, more user-friendly, less expensive, more portable, and more powerful. A discussion of some of the key advantages of computer-based assessment in neuropsychology follows.

One advantage of computer-administered assessments is the ability to assess several aspects of human behavior simultaneously. A computer can

record and score responses, present multiple stimuli simultaneously, and time these simultaneous events all at the same time. Commenting on this advantage, Kane and Kay (1997) wrote, "The ability of the computer to assess divided attention is probably one of the most significant contributions of computer-based testing to the assessment process" (p. 361). Measures of divided attention are especially sensitive to mild brain dysfunction and the computer provides the ability to evaluate the domains of attention and executive functioning simultaneously (Kane & Kay). Therefore, computer programs provide a way of assessing cognitive processes that were next to impossible for human test administrators to assess.

Another advantage of computer-administered tests is the ability to exercise accurate control over the presentation of stimuli (Kane & Kay, 1997) Computers have the ability to adjust screen graphics for visually impaired persons, produce numerous sounds at different intensities without distortions, manage presentation time (Golden, 1987), and have visual displays that move (Pellegrino, Hunt, Abate, & Farr, 1987). Along with the ability to manipulate stimuli, programs can also present stimuli in a predetermined order, in random order, or in an order based on the examinee's performance (Adams & Heaton, 1987; Golden). Computer programs allow for a level of control over stimuli that could never be achieved through traditional test administration.

Another aspect of assessment that is unattainable by human beings is the ability of computers to provide exact timing. In neuropsychological

assessment, many tests that are sensitive to brain damage have speed scores (Kane & Kay, 1997). The ability of computer programs to easily and unobtrusively provide accurate timing of multiple functions allows psychologists to assess aspects of behavior that, in the past, have not been measured easily, such as response latencies and overall performance time (Adams & Heaton, 1987; Long & Wagner, 1986). In addition to response latencies, the computer can measure other aspects of performance quality such as force and motor steadiness (Adams & Heaton, 1987). Thus, computer programs provide more accurate measuring procedures and allow for the assessment of a variety of behaviors that human administrators can not easily measure.

Computers are also excellent administrators and accountants, excelling in accuracy, recording, and speed while at the same time being cost effective (Butcher, Perry, & Atlis, 2000; Collie, Darby, & Maruff, 2001; Elwood, 1969; Kane & Kay, 1997). Computers can relieve clinicians of the routine tasks related to scoring, converting raw scores to standardized scores, and making normative comparisons. This efficiency with administrative tasks can assist neuropsychologists with getting reports out in a timely manner (Butcher, 1987). Another time saver comes in the form of the speed. Research has shown that many computer-administered tests take half the time of the paper-and-pencil versions of the same test (Watts, Baddeley, & Williams, 1982). This time reduction translates into cost reduction. As discovered by Elwood (1972c), the

cost of administering and scoring an automated version of a test was 50% lower than the face to face version. The efficiency of the computer-based assessment could help address the increasing need for less expensive neuropsychological services (Space, 1981) as well as free up clinicians' time for tasks that require their expertise.

As the administrator of a test, the computer is free from examiner bias (Butcher, 1987; Kane & Kay, 1997). The computer does not provide unintended cues as any human examiner will. Computer programs provide objective interpretations that are not influenced by subjective interpreters, like clinically generated conclusions. In particular, in some sensitive social situations, Smith (1963) along with others (Long & Wagner, 1986) believed that computer-individual interaction may be less stressful and less biased than face-to-face interactions. Johnson and Mihal (1973) agreed with these conclusions. In their study, African-Americans performed worse than Anglo-Americans on the human administered version of an academic abilities test. However, on the computer-administrated form there was no significant difference between the two races. This is a profound finding and a strong argument for the potential fairness of computer-based testing. As noted by Johnson and Mihal, the lack of human biases may increase reliability and validity of the test.

The consistent, non-biased administration of computerized neuropsychological assessments contributes to a standardization of procedure unattainable with human examiners. Improved standardization increases

reliability through standard presentation and scoring (Elwood, 1969; Long & Wagner, 1986). Standardization is also easy to maintain across various testing centers. Since computers are not susceptible to "off days" like human administrators (Butcher, 1987), consistent adherence to testing protocol is easily maintained.

Another advantage of computer-administered tests is that they are excellent tools for adaptive testing, a more complex administration of test items that is similar to basal and ceiling rules on paper-and-pencil tests (Weiss & Vale, 1987; Wetzler, 1990). Adaptive testing is a form of testing that allows for shorter tests by determining the appropriate starting places on a test, based on the performance on prior items. During adaptive testing the computer "presents the items to the examinee, receives and scores the item responses, chooses the next item to administer, and terminates the test when appropriate" (Weiss & Vale, p. 326). The computer has the ability to administer individualized tests to each examinee in an efficient and accurate manner. "In essence, the computer determines a complex branching strategy. Paper-and-pencil administration does not offer this flexibility" (Wetzler, p. 54). Adaptive testing improves efficiency and saves time for the examinee (Roid, 1986).

The many advantages of computer-based programs make computer programs excellent for repeated testing scenarios. First, the computer has the capacity to generate multiple forms of a test through storage of a large library of stimuli and questions, or through programs that generate new test items (Adams

& Heaton, 1987; Kane & Kay, 1997; Roid, 1986). Multiple equivalent forms of an assessment are essential in repeated testing scenarios. Second, computer-based assessments can be administered to several individuals at the same time, making it ideal for sports programs that need baseline data on their athletes. Third, computer-based tests are also portable, which facilitates testing individuals in their own environment (Collie, Darby, & Maruff, 2001; Schatz & Browndyke, 2002).

Finally, most research has demonstrated equivalency of computer-based tests with their traditional counterparts (Campbell et al., 1999; Choca & Morris, 1992; Elwood, 1969, 1972b, 1972d; French & Beaumont, 1990; Knights, Richardson, & McNarry, 1973; Pelligrino, Hunt, Abate, & Farr, 1987; Roid, 1986; Schatz & Browndyke, 2002; Watts, Baddeley, Williams, 1982; Wingenfeld, Holdwick, Davis, & Hunter, 1999). Equivalency means that scores from the computer programs have demonstrated high correlations with scores from their paper-and-pencil counterparts. Many of these programs have also demonstrated that they are measuring the same constructs as the traditional test through factor analysis and construct validity research. Research on computer-based neuropsychological test batteries is also demonstrating that these tests have the sensitivity and specificity necessary to be considered legitimate clinical tools (Fray, Robbins, & Sahakin, 1996; Green, Green, Harrison, & Kutner, 1994; Ritchie & Hallerman, 1989; White et al., 2003). In fact, some researchers are finding that computer-based neuropsychological assessment is

more sensitive than many traditional neuropsychological measures (Fowler, Saling, Conway, Semple, & Louis, 1997). With all of the convincing advantages of computerized neuropsychological assessment, there still remain some concerns about using this type of testing.

Limitations of Computer-based Neuropsychological Assessment

The clinician's use of computers in neuropsychological evaluations has lagged behind the development of computer-based neuropsychological assessment. Clinicians appear to be hesitant to incorporate the computer as a part of the assessment process (Kane & Kay, 1997). This hesitancy can be attributed to some of the valid concerns over the limitations associated with computerized assessment. Some of the main limitations of computer-based assessment are discussed in this section.

Long and Wagner (1986) stated "that...attempts to convert a test for computer application may radically change what the test measures and invalidate available norms" (p. 553). Agreeing with this view, Hoffer and Green (1985) suggested that factors tangential to the administration of the test may negatively impact an individual's test performance and therefore computer-based tests may not really be equivalent to paper-and-pencil versions. There are some studies that suggest factors external to the actual test might contribute to the variance found in computer-administered tests (Browndyke et al., 2002; Johnson & White, 1980; Lee, 1986). For example, Browndyke et al. found that as computer-related anxiety increased, performance on the

computerized version of the Category Test decreased. Because the sample size in the study was small and the participants only received the computerized version of the test, questions regarding other factors that may have influenced participants' performance on the test remain; e.g., how does test-related anxiety impact computer related-anxiety and performance? In spite of the shortcomings of the study, results might suggest that some computerized assessments could be measuring different factors than paper-and-pencil tests. In looking at other aspects that might impact computer-based assessment, Lee found that past computer experience did account for a significant amount of variance on a computerized arithmetic reasoning test. To address this issue, researchers have found that computer training positively impacts scores on a computer-based assessment (Johnson & White). More research is needed in this area to determine what additional factors might contribute to the outcomes on computer-based tests.

Another concern with computer-administered tests is that some of the software on the market is poorly designed. While excellent timing procedures are available, there are several programs that still are inaccurate in this regard. For example, programs that use keyboards as input devices are particularly susceptible to timing inaccuracies. However, "A computer at its worst is a more accurate and reliable timekeeper than the psychologist holding a stopwatch in one hand and attempting to manipulate test stimuli with the other" (Kane & Kay, 1997, p. 367). Another concern about some of the software on the market

it may be difficult to make the necessary adjustments to assess people with special needs. For example, computers may not be appropriate for some individuals with handicaps, peripheral injuries, sensory loss, or language deficits (Golden, 1987). Another problem with less contact between the examinee and examiner is the lack of qualitative behavioral evaluation, which is so essential to neuropsychological assessment. A computer cannot assess motivation, fatigue, resistance, variations in levels of consciousness, stress, coping style, and response to feedback, which are all helpful tools in neuropsychological assessment (Long & Wagner). However, clinicians could use computerized assessment to free up their time to observe these behaviors while not being distracted with test administration responsibilities.

The proliferation and availability of new computer-based neuropsychological assessments adds another problem. It is now easy to exploit this new technology (Butcher, 1987; Kane & Kay, 1997). Untrained individuals have access to computer-based assessments and the results of those assessments. People may believe that because the results came from a computer, they are more valid than a report from a clinician. Long and Wagner (1986) expressed concern that this exploitation could manifest itself in untrained professionals using these tests and basing life-changing decisions on an automated test, without consulting professionals.

With all of these limitations, clinicians are cautious about using computers. Continued caution and education about computerized assessment,

or any assessment for that matter, is appropriate. As with paper-and-pencil tests, clinicians need to be careful about issues of reliability, validity, norming data, and whether or not computerized assessments are thoroughly researched (Kane & Kay, 1997). The principles that apply to paper-and-pencil test development are also applicable to computerized test development (Kane & Kay, 1992). To continue to address the issues of reliability and validity, Schatz and Browndyke (2002) recommended "the development of psychometric databases on comparisons with long standing empirically sound test measures; [and] additional validation of measures by parties not involved in their commercial development" (p. 395). Other researchers agree, saying that, "The need for large-scale validation studies of computerized versions of tests in clinical conditions cannot be stressed too strongly" (French & Beaumont, 1990).

While the limitations of computerized assessment should cause clinicians to be thoughtful about how these assessments are used, neuropsychologists should no longer ignore the advances and advantages that this technology brings to the neuropsychological field. Computerized assessment is fast, cost-effective, accurate, and unbiased. The computer can assess aspects of human behavior that are unattainable with human administrators. There are computer programs that have been and continue to be thoroughly researched, with information about reliability, validity, factor structure, and norms. In some cases computer-administered neuropsychological

tests may be the best choice for evaluating specific cognitive concerns. It is nearing the time when neuropsychologists need to consider responsibly incorporating computer-based assessments as valid tools in neuropsychological evaluations.

Introduction to the MicroCog

A significant portion of the information in this section was obtained from the MicroCog: Assessment of Cognitive Functioning Manual (Powell et al., 1993).

One of the neuropsychological computer-administered assessments that is normed and standardized for clinical implementation is the MicroCog: Assessment of Cognitive Functioning (Powell, et al., 1993). The MicroCog is a commercially available neuropsychological test of cognitive functions that screens for mild to moderate levels of cognitive impairment (Elwood, 2001). A medical insurance carrier, The Risk Management Foundation of the Harvard Medical Institutions, funded the development of the MicroCog (Powell et al.). The instrument was first designed to screen elderly physicians and other professionals for subtle changes in cognitive functioning, in hopes of reducing malpractice liability (Elwood). Originally called Assessment of Cognitive Skills (ACS), the instrument was purchased by the Psychological Corporation and was renamed, MicroCog (Powell et al.). The new instrument had the number of items reduced, incorporated response-time in scoring (Douglas

Powell, personal communication, November 4, 2003), and was normed as a general neuropsychological screen (Kane, 1998).

The MicroCog was developed with four goals in mind (Powell et al., 1993). The first was to use a theoretical model of brain-behavior relationships, relating basic domains of behavior and their underlying brain mechanism, to develop the content. Second, the authors wanted to incorporate measures of information processing speed that would be clinically meaningful. Third, the interaction of education and performance on cognitive measures was considered by creating measures with extremely high ceilings and creating norms based on level of education. Finally, the developers of the MicroCog wanted to provide norming information into the ninth decade to address the changing demographics in the elderly population (Powell et al.).

The MicroCog is available in two versions; the Short Form, which includes 12 subtests with a completion time of about 30 minutes, and the Standard Form, which is comprised of 18 subtests and requires about 60 minutes to complete (Powell et al., 1993). A weakness of the MicroCog is that there are no alternate forms available. A recent change in the newest version of the test released in 1996 does allow for the examiner to select specific tests for each examinee (Elwood, 2001). For the purposes of this study, only the characteristics of the Standard Form will be discussed.

Scores from the 18 subtests of the MicroCog contribute to three levels of index scores (see Figure 1). The first level of index scores is made up of five

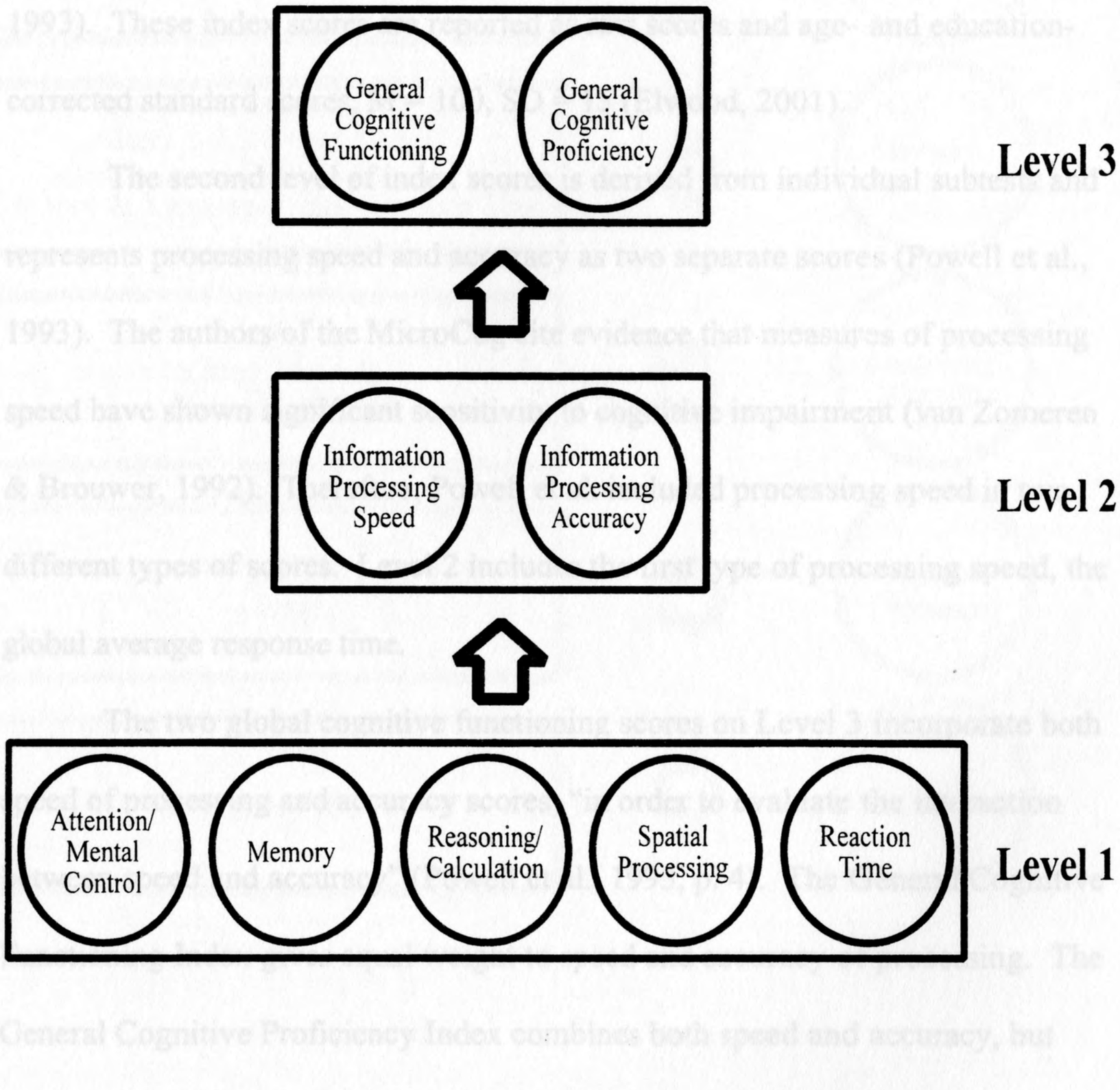


Figure 1: Levels of MicroCog Index Scores

Note. From MicroCog: Assessment of Cognitive Functioning (p. 1), by Powell et al., 1993, San Antonio, TX: Harcourt Assessment, Inc. Copyright 1993 by the Risk Management foundation of the Harvard Medical Institutions, Inc.

neuropsychological content domains: Attention/Mental Control, Memory, Reasoning/Calculation, Spatial Processing, and Reaction Time (Powell et al., 1993). These index scores are reported as raw scores and age- and education-corrected standard scores, M = 100, SD = 15 (Elwood, 2001).

The second level of index scores is derived from individual subtests and represents processing speed and accuracy as two separate scores (Powell et al., 1993). The authors of the MicroCog cite evidence that measures of processing speed have shown significant sensitivity to cognitive impairment (van Zomeren & Brouwer, 1992). Therefore, Powell et al. included processing speed in two different types of scores. Level 2 includes the first type of processing speed, the global average response time.

The two global cognitive functioning scores on Level 3 incorporate both speed of processing and accuracy scores, "in order to evaluate the interaction between speed and accuracy" (Powell et al., 1993, p. 4). The General Cognitive Functioning Index gives equal weight to speed and accuracy of processing. The General Cognitive Proficiency Index combines both speed and accuracy, but gives greater weight to the accuracy score.

Content of Level One Domains

On Level 1, each domain is comprised of several subtests that were designed to assess the cognitive components described by that domain (See Figure 2). A description of what each domain measures and the subtests that

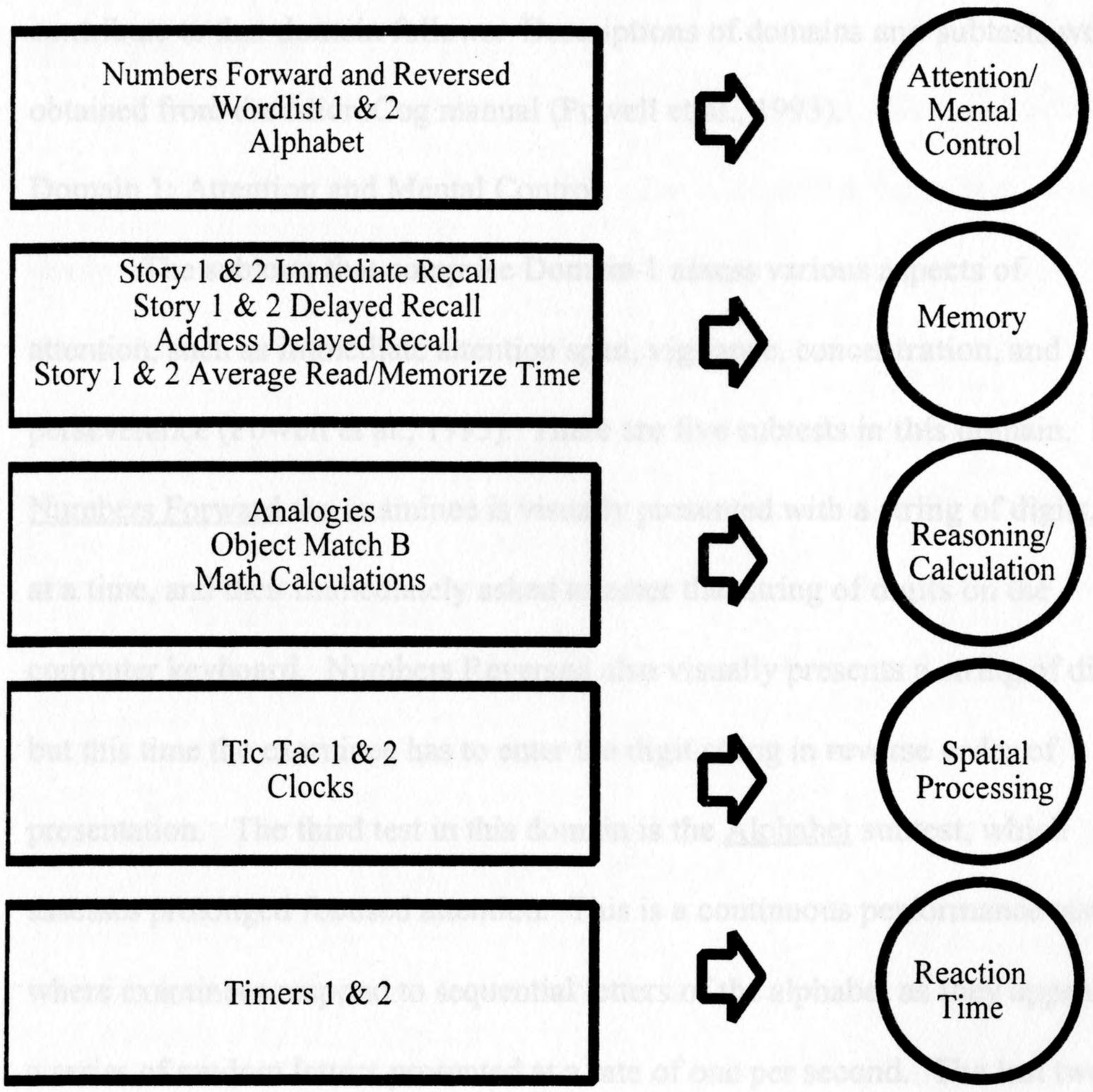


Figure 2: Subtests Contributing to Level 1 Index Scores

Note. From MicroCog: Assessment of Cognitive Functioning (p. 1), by Powell et al., 1993, San Antonio, TX: Harcourt Assessment, Inc. Copyright 1993 by the Risk Management foundation of the Harvard Medical Institutions, Inc.

contribute to that domain follows. Descriptions of domains and subtests were obtained from the MicroCog manual (Powell et al., 1993).

Domain 1: Attention and Mental Control

The subtests that comprise Domain 1 assess various aspects of attention, such as immediate attention span, vigilance, concentration, and perseverance (Powell et al., 1993). There are five subtests in this domain. For Numbers Forward the examinee is visually presented with a string of digits, one at a time, and then immediately asked to enter that string of digits on the computer keyboard. Numbers Reversed also visually presents a string of digits, but this time the examinee has to enter the digit string in reverse order of presentation. The third test in this domain is the Alphabet subtest, which assesses prolonged focused attention. This is a continuous performance task where examinees respond to sequential letters of the alphabet as they appear in a series of random letters presented at a rate of one per second. The last two subtests in this domain, Wordlist (1 and 2) assess attention, vigilance, mental control, and incidental learning (Powell et al.). For Wordlist 1, the examinee is asked to view words one at a time and respond to only the words that fit a category provided by the instructions. The second part of Wordlist involves incidental learning, where the examinee is asked to respond only to the words that were in Wordlist 1.

Domain 2: Memory

Domain 2 consists of six subtests that measure immediate and delayed recognition memory (Powell et al., 1993). The contents of the subtests were designed to assess the kinds of memory problems experienced by an older population (Powell et al.). Story 1, Immediate Recall, asks the examinee to read a brief story and then immediately respond to multiple choice questions about the story. Story 1, Delayed Recall occurs approximately 20-30 minutes after the original story was read, and again asks the examinee to respond to multiple choice questions about Story 1. Procedures for Story 2, Immediate Recall, and Story 2, Delayed Recall are identical to the procedures described for Story 1, using a different narrative. For Address, the computer visually displays a name and address and asks the examinee to memorize the information. Later, for Address, Delayed Recall the examinee is asked to respond to multiple-choice questions concerning the person's title, name, and address (Powell et al.).

Domain 3: Reasoning and Calculation

In this domain, which consists of three subtests, the authors wanted to assess abstraction and reasoning, because impairment in these areas is highly associated with brain injury (Powell et al., 1993). The Analogies subtest measures inductive reasoning. For this task the examinee is given a list of three words, of which the first two have a particular relationship. Then, the examinee is asked to select a word from a list of words that will best express that same

relationship with the third word (Powell et al.). Assessing concept formation and cognitive flexibility, the Object Match subtest presents the examinee with four figures, and then asks the examinee to select one figure that does not match the other three. The examinee is then asked to do the same process again with the same four figures using a different grouping. The Math Calculations subtest was designed to measure the ability to perform arithmetic operations. However, due to the awkward entry of the final answer, this has turned out also to be a test assessing working memory (Devivo, Rothlan, Price, & Fein, 1997; Powell, 1997). Math Calculations presents addition, subtraction, and multiplication, and division problems. Examinees are asked to solve these problems, taking as long as they need (Powell et al.). Then they are to enter the answers using the numeric keypad, from left to right (this is different from solving an arithmetic problem on paper where the answer is solved from right to left).

Domain 4: Spatial Processing

The Spatial Processing domain, consisting of two subtests, assesses both novel and familiar visual spatial processing and memory (Powell et al., 1993). The Tic Tac 1 and 2 subtests assess short-term recall of novel geometric patterns. This task displays a pattern of darkened squares on a 3X3 matrix for a short time (Powell et al.). The examinee must then reproduce the pattern by selecting the appropriate keys on the 3X3 numeric keypad. The second subtest, Clocks, assesses visuospatial analysis of familiar stimuli (Powell et al.). This test displays seven clock faces with hour and minute hands, but no numerical

markers. The examinee is then asked to select the correct time from a list of possible times.

Domain 5: Reaction Time

This last domain has two subtests and was designed to assess simple reaction time in both auditory and visual modalities. Timers 1 is presented at the beginning of the MicroCog and Timers 2 is administered at the end of the testing session. Both of these tests measure the speed with which an individual responds to visual and auditory signals (Powell et al., 1993).

The authors of the MicroCog included two administrations of some subtests. This was done to determine whether individuals tired over the course of the test, resulting in a decline in performance. The authors also wanted to understand the effects of delay and interference on verbal memory (Powell et al., 1993).

Design of the Human-Computer Interface

Powell et al. (1993) designed the operation of the MicroCog to be as simple as possible. This was done to reduce any potential negative impact on examinees who have no or little prior computer experience. While the keyboard is used as an input device, only the Numeric keypad (0-9), the Backspace and Enter keys, and the letter P (to pause the program) are used during the test. The pause function was included in the test design to allow examinees to take a break from testing as they choose. It was hoped that this function would help examinees feel more in control of the process and less pressured by time

(Powell et al.). "The test begins with a brief introduction to the keyboard and reviews the keys necessary for completion of the test. The interface appears to be manageable for the majority of patients likely to be seen in clinical settings for screening" (Kane & Reeves, 1997, p. 435).

Each subtest on the MicroCog is preceded by a set of instructions displayed on the computer screen. While the examinee is not allowed trial items, many instruction sets provide examples of the types of items that will be answered (Powell et al., 1993). To reduce frustration, consecutive errors can result in the termination of certain subtests. Powell et al. recommended that the examiner provide a brief introduction to the test, and then remain available for questions or problems.

Standardization and Norms

The standardization sample for the MicroCog was selected to be representative of the U.S. population, using the 1988 Census (Powell et al., 1993). The sample consisted of 810 adults from the ages of 18 to 89 placed in nine age groups (18-24, 25-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, and 80-89). There were an equal number of males and females in each group, and three racial groups (African American, Hispanic, and Anglo American) were sampled in a ratio that represented the U.S. Census. The three educational levels that were sampled were less than high school, high school, and greater than high school (Powell et al.). The average number of years of schooling for the group that had education beyond high school was 15.6 years and the median

was 16. Since physicians have at least 22 years of education, the educational levels may not provide a valid comparison group for physicians.

Kay and Kane (1997) noted that the MicroCog (Powell et al., 1993) norms were developed using data from two versions of software, an older version and a newer version. The older version of MicroCog software used a timing procedure that used the system clock and was infamously inaccurate. The newer software contained a more sophisticated timing procedure that had only minor timing inaccuracies. Of the 810 individuals in the sample, 500 were tested with the older, less accurate software, and 310 were tested with the newer, more accurate software (Kay & Kane). Therefore, clinicians need to consider these inaccuracies when comparing an individual's performance on the MicroCog to the published norms.

The MicroCog manual (Powell et al., 1993) also provides data on a variety of special groups. These groups of people include those with Dementia, Depression, Lobectomy, Lupus, Alcoholism, Schizophrenia, Psychiatric/Neurologic problems, and Physicians. The information on physicians does not include key demographic variables such as age. The authors warn that data on the special groups are not representative samples of these populations, and are intended only as exemplars (Powell et al.). Therefore, there are no norms for physicians with the current version of the MicroCog.

Reliability

The MicroCog manual (Powell et al., 1993) defines reliability as "the consistency, precision, and stability of test scores across situations" (p. 63). Reliability coefficients were calculated for individual subtests and for the nine index scores. For most subtests, internal consistency was assessed using the split-half methodology and the two halves were correlated and adjusted for the score on the complete subtest using the Spearman-Brown formula (Powell et al.). "Otherwise, generalizability coefficients were employed" (Kane & Kay, 1997, p. 374) for Numbers Forward, Numbers Reversed, Address, Story 1 & 2, and Tic Tac. "The average subtest total score and response time reliability coefficients... ranged from .58 to .90, with a mean of .76. Most are at or above .70, which is generally considered minimally acceptable for individual subtest scores" (Elwood, 2001, p. 91). The nine index scores had reliability coefficients ranging from .83 to .95, well within acceptable limits (Kane & Kay).

Test-retest stability was assessed using 262 adults who were tested on two separate occasions, with a range of 27 to 255 day intervals (Powell et al., 1993). For the individual subtests and nine indexes, test-retest reliability was presented in the form of means and standard deviations (Kane & Kay, 1997). The authors noted that, "the means from test to retest are very stable and show little practice effects" (Powell et al., p. 73). The "decisions consistency" reliability coefficients reflect the consistency of classification into one of four

categories (below average, low average, average, and above average). The decisions consistency reliability coefficients were reasonably stable over time (Powell et al.). The nine index decisions consistency reliability coefficients were higher (ranging from .73 to .99) than the individual subtest coefficients (ranging from .64 to .96 for average response time and total score).

In his comprehensive article on the MicroCog, Elwood (2001) provided his conclusion regarding standardization and reliability of the MicroCog. He stated, "Overall, MicroCog was normed better than most neuropsychological tests when it was introduced in 1993 and, despite minor criticisms, still compares favorably with current neuropsychological test batteries" (p. 92). While the MicroCog is a well normed test for the general population, there are no current norms for physicians.

Validity

To determine validity for an instrument, research needs to demonstrate that the instrument measures what it purports to measure (Anastasi, 1988). To demonstrate validity, Powell et al. (1993) consulted experts, calculated intercorrelations of scores, conducted factor analyses, conducted comparison studies and classification studies, and researched performance in special groups. Along with the authors' works, a few independent studies have been done to assess various validity issues. A brief discussion of these validity activities follows.

Content Validity

One of the original goals for creating the MicroCog was to base the selection of content on a theoretical foundation of brain-behavior relationships (Powell et al., 1993). The authors asked neuropsychologists, neurologists, and psychiatrists to examine the MicroCog for appropriate content coverage. Several pilot studies were conducted to assess aspects of the test, including content. Powell et al. incorporated feedback from these consultants and from pilot sessions, and made pertinent revisions.

Construct Validity

The intercorrelations of the subtests and indexes confirmed the expected relationships between the subtests and nine index scores (Powell et al., 1993). In other words, the subtests that comprised a particular domain were more highly correlated with that domain index score than with other domain index scores. Numbers forward and backward showed a fairly high correlation (Powell et al.). The General Cognitive Functioning correlations between the Standard Form and the Short Form were very high, ranging from .95 to .97.

Factor Analysis

An exploratory factor analysis resulted in two factors that were present through all of the age groups (Powell et al., 1993). The first factor, labeled Information Processing Accuracy, appears to include overall cognitive abilities. The second factor is Information Processing Speed. This factor has been shown to be highly sensitive to brain injury (van Zomeren & Brouwer, 1992).

Criterion-Related Validity

Powell et al. (1993) compared performance on the MicroCog between various clinical groups and non-clinical groups. The groups were matched with respect to age, education, gender, and ethnicity. Statistics for sensitivity, specificity, and whether or not groups were correctly classified were calculated. The average classification rate for two different dementia groups was 92%. The classification rates for the Schizophrenia and Psychiatric/Neurologic groups were between 85% and 88%. A low classification rate in the Depression group (63%) was perhaps indicative of a low sensitivity to depression (Powell et al.). Specificity and sensitivity rates were above 80% for both of the Dementia groups, the Schizophrenia group, and the Psychiatric/Neurologic group.

Green, Green, Harrison, and Kutner (1994) conducted "the first major independent study of the MicroCog" (Elwood, 2001). They compared 52 participants with mild cognitive impairment (mean age, 71.2) and 50 unimpaired (mean age, 68.7), age- and education- matched control participants. Using a global cut-off score that assessed only accuracy and not processing speed (Elwood, 2001), the MicroCog (then ACS) demonstrated a sensitivity of .83 and specificity of .96. Green et al. concluded that the MicroCog was a valid screening instrument for detection of mild cognitive impairment in an elderly population.

Lopez (1999) conducted a study with 346 participants who were substance abusers. He found that the Short Category Test (Wetzel & Boll,

1987) was correlated with the Attention and Reasoning domains. The MicroCog (Powell et al., 1993) found group differences between the standardization sample and the substance abuse group, with a classification accuracy of 76%.

Nelson (1999) studied the MicroCog as an outcome measure of cognitive improvement made after traumatic brain injury. She determined that the MicroCog (Powell et al., 1993) was sensitive to cognitive changes made by participants in a rehabilitation program. She also found that expert rating of change was commensurate with the change measured by the MicroCog.

Convergent and Discriminant Validity Studies

This section examines the relationship between four of the five cognitive domains and related external criterion measures, as reported in Powell et al. (1993). The fifth domain, Reaction Time, was not included in the concurrent studies. High correlations with similar measures give credence to convergent validity. Low correlations with measures of different constructs provide evidence of divergent (discriminant) validity. In other words, divergent validity is demonstrated when the external criterion measures for one domain have lower correlations with the other three domains (Powell et al.).

Domain 1: Attention and Mental Control. Convergent validity was demonstrated with the following external criterion measures: Trail Making A and B (Reitan, 1958), Attention and Initiation/Perseveration scales from the Dementia Rating Scale (DRS; Mattis, 1973), and Attention/Concentration Index

and Digit Span Forward and Backward from the WMS-R (Wechsler, 1987). All of the correlations between the MicroCog - Attention and Mental Control domain and the external criterion measures were significant, with correlations ranging from .36 (Digit Span Forward, WMS-R) to .57 (Attention/Concentration Index, WMS-R; Powell et al., 1993).

Domain 2: Memory. Convergent validity was demonstrated with the following external criterion measures: Rey Auditory-Verbal Learning Test (Rey AVLT; Rey, 1964), the Memory subtest from the DRS (Mattis, 1973), and the four Memory indexes from the WMS-R (Wechsler, 1987). All of the correlations between the Memory Index of the MicroCog and the external criterion measures mentioned above were significant and ranged from .28 (Rey AVLT, Trial 1) to .46 (Delayed Recall Index, WMS-R; Powell et al., 1993).

Domain 3: Reasoning and Calculation. Convergent validity was demonstrated with the following external measures of reasoning and calculation: Arithmetic, Similarities, and Block Design subtests of the WAIS-R (Wechsler, 1981), the Conceptualization and Initiation/Perseveration subtests from the DRS (Mattis, 1973), and the Abstraction Quotient and Abstraction T-Score from the Shipley Institute of Living Scale (Shipley, 1946; Zachary, 1992). All of the correlations were significant and ranged from .28 (Shipley's Abstraction Quotient) to .56 (Shipley's Abstraction T Score; Powell et al., 1993).

Domain 4: Spatial Processing. Convergent validity was demonstrated with the following external criterion measures: Rey-Osterrieth Complex Figure Test (Rey-Osterrieth CFT; Rey, 1941, 1959; direct copy), Visual Reproduction I and II and the Visual memory Index from the WMS-R (Wechsler, 1987), the Construction Index from the DRS (Mattis, 1973), and the Benton Visual Form Discrimination Test (Benton, Hamsher, Varney, & Spreen, 1983). All of the correlations were significant and ranged from .26 (Benton VFDT) to .37 (Construction Index, DRS; Powell et al., 1993).

Green, Green, Harrison, and Kutner (1994) had similar results. They found that the MicroCog's (then ACS) subtests correlated significantly with traditional neuropsychological tests measuring similar cognitive domains. "However, the convergent-divergent construct was not supported, in that most of the ACS subtests were related at least as strongly to other less similar CNTs [conventional neuropsychological tests] as they were to the CNT designated most similar" (Green et al., p. 783). Therefore, Green et al. recommended caution when using individual subtests to make inferences about an examinee.

Another study (O'Keefe, 1997) found that the four domains of the MicroCog (Powell et al., 1993) demonstrated convergent validity with well-established neuropsychological tests that measured similar cognitive domains. However, some domains were better supported than others, and again caution was recommended when using individual subtests to draw conclusions about an examinee.

In summary, the MicroCog (Powell et al., 1993) is a computer-administered screening assessment that has the capacity to detect cognitive impairment in an older population (Green et al., 1994). However, its ability to detect cognitive decline at other ages, or to differentiate between dementia and other mental disorders has not been well established (Elwood, 2001). There is a need for continued research to better understand the limits and strengths of the MicroCog. Having been developed for physicians, "Microcog shows virtually no ceiling effect and thus can detect cognitive deficits in well-educated, higher functioning individuals" (Elwood, p. 99). The MicroCog was designed as a screening device and was not intended to replace a full neuropsychological evaluation. However, as a screening instrument, the MicroCog provides clinicians with a valid, reliable, and well normed assessment tool (Elwood).

Implications of the Literature for the Current Study

In conclusion, the technology and software for computerized neuropsychological assessment have developed to the point that some neuropsychological screening instruments have the sensitivity and specificity to be confidently used by neuropsychologists (Kane & Kay, 1992). Some screening instruments appear to have the ability to detect neurological deficits that are undetectable by traditional methods (Adams & Heaton, 1987; Long & Wagner, 1986). As always, reliability and validity are key considerations in selecting neuropsychological instruments. Therefore, psychometrically sound

computerized neuropsychological instruments are valid neuropsychological tools.

One instrument that was developed with particular attention to reliability and validity concerns was the MicroCog: Assessment of Cognitive Functioning (Powell et al., 1993). This test was originally developed to detect mild cognitive deficits in aging physicians. The validity of the index scores and global functioning scores is well within acceptable ranges. The reliability of the MicroCog has also been demonstrated to be acceptable (Powell et al.). Since this test was originally designed for physicians or high functioning professionals, most subtests do not have ceilings (Powell et al.). Therefore, individuals with high levels of education can still obtain feedback on levels of, and changes in, cognitive functioning. Thus, this test is ideal as a neuropsychological screen for physicians involved in competency evaluations.

One of the few weaknesses of this assessment is the lack of norms for physicians. The original creators of this instrument, at that time called Assessment of Cognitive Skills, developed norms for physicians (Powell & Whitla, 1994). However, the test was sold to Psychological Corporation, which made changes in number of items and timing methods, and did not norm the new version on physicians (Douglas Powell, personal communication, November 4, 2003). Therefore, information about how a normal population of physicians performs on this new test would be helpful in evaluating physicians

who take the MicroCog (Powell et al., 1993) as a part of a competency evaluation.

The inclusion of neuropsychological assessment as an essential component of physicians' competency evaluations is a fairly new occurrence. Neuropsychological assessment needs to remain a part of physician competency evaluation programs. First, physicians are vulnerable to neuropsychological impairment as a result of substance abuse, traumatic brain injury, disease, and normal aging (Madden, 1988). Second, research has demonstrated that there is a relationship between physicians' neuropsychological performance and clinical performance. Thus, neuropsychological impairment can adversely impact a physician's ability to competently practice medicine (Reich et al., 1999; Schueneman, Pickleman, Hesslein, & Freeark, 1984). Finally and most importantly, early detection of neuropsychological impairment may lead to treatment to reverse the effects of problems resulting in impairment, or a respectful evidence-based decision to limit or terminate a physician's practice (Kapur, 1997). In both cases the physician is protected and the medical profession maintains the public trust.

Little is known about the neuropsychological difference between physicians referred for competency evaluation and physicians who are practicing successfully (Thompson, 2003). A few studies suggest that there might be major differences in the cognitive constellation of the two groups (Thompson; Williams, Williams, & Norcross, 2002). However, due to either a

lack of control groups or in one study a very small control group, there is a need for research projects that incorporate sufficiently large control groups. The addition of control groups will provide empirically-based support for there being a difference in the cognitive constellation of physicians referred for competency evaluations and physicians not involved in such an evaluation.

The purpose of this study was to provide more information about the cognitive differences between physicians referred for competency evaluations and physicians who were practicing successfully.

CHAPTER III

METHODS

Introduction

This chapter includes a discussion of the methodology for this study and is comprised of a description of the participants in the database obtained from the Center for Personalized Education for Physicians (CPEP), an overview of the strategies for recruiting participants, a discussion of the measures that were employed, a description of the procedures that were used, and the hypotheses and statistical analyses that were implemented.

Participants

An existing database of 361 physicians who completed competency evaluations from January 1997 to January 2004 was obtained from CPEP. These physicians took the MicroCog (Powell et al., 1993) as a part of an intensive two to three day evaluation conducted by CPEP. The MicroCog (Powell et al.) was administered in a small testing room on a desktop computer with an examiner present to answer any questions the examinee might have. As stated in Chapter Two, physicians were referred for evaluation for a variety of reasons, such as concerns about clinical competence; concerns about professional or interpersonal behavior; documentation problems; prescribing problems; renewing, restoring and/or seeking licensure or privileges; patient

complaints; illness or injury; and other reasons (Stulp et al., 2004). Physicians who were foreign born, foreign trained or educated, whose primary language was not English (Jacobs et al., 1997), who took the MicroCog (Powell et al.) more than once, and for whom there had been no previous quality of care concerns, were all eliminated from the database. Evaluations for Physician Assistants were also removed from the database. A total of 94 (26%) physicians were excluded from the CPEP database, leaving 267 physicians. The age range of this group was from 31 years to 76 years, with a mean age of 51.5 years and standard deviation of 9.11. Physician specialties were grouped into three categories: 47.2% Primary Care Physicians, 36.7% Surgery Specialties, 16.1% Other Specialties. Please see Table 3 for further details about this physician group. Also, please see Appendix A for University of Denver Institutional Review Board approval and Appendix B for the HIPPA Waiver of Authorization Form.

For the second data set, 68 physicians were recruited from the greater Denver area to participate in this study. Recruitment strategies will be discussed in the Procedure section of this chapter. The age range of this group was from 30 years to 80 years, with a mean age of 49.4 years and standard deviation of 12.5. As with the data from CPEP, physician specialties were grouped into three categories: 42.6% Primary Care Physicians, 23.5% Surgery Specialties, and 33.9% Other Specialties. The average years of practicing medicine after residency was 19.2 years, with a standard deviation of 12.6

Table 3 Demographics for Two Physician Groups

	Control	Competency
N	68	267
Gender		
Male	41 (60.3%)	223 (83.5%)
Female	27 (39.7%)	44 (16.5%)
Handedness		
Left	7 (10.3%)	24 (9%)
Right	61 (89.7%)	238 (89.1%)
Ambidextrous	0	2 (.7%)
Not Specified	0	3 (1.1%)
Specialty		
Primary Care	29 (42.6%)	126 (47.2%)
Surgical	16 (23.5%)	98 (36.7%)
Other	23 (33.9%)	43 (16.1%)
Age		
M	49.4	51.5
SD	12.5	9.1

study made a prediction about the comparison of the proportion of physicians with neuropsychological impairment in the two physician groups. Eliminating control group physicians with neuropsychological concerns would have

artificially impacted that result. Also, physicians with neuropsychological related illnesses or insults were included in the competency evaluation group.

Measures

Two measures were employed in this investigation: the MicroCog: Assessment of Cognitive Functioning (Powell et al., 1993) and a demographic questionnaire developed specifically for this study. A description of each instrument follows.

MicroCog: Assessment of Cognitive Functioning (Powell et al., 1993)

The purpose of this instrument was to assess neurocognitive functioning of physicians. For a more complete description of the MicroCog (Powell et al., 1993), please see Chapter Two, pages 61 - 79.

The MicroCog is a commercially available neuropsychological test of cognitive functioning which screens for mild to moderate levels of cognitive impairment (Powell et al., 1993). A medical insurance carrier, The Risk Management Foundation of the Harvard Medical Institutions, funded the development of the MicroCog (Powell et al.). The instrument was first designed to screen elderly physicians and other professionals for subtle changes in cognitive functioning in hopes of reducing malpractice liability (Elwood, 2001). Originally called Assessment of Cognitive Skills (ACS), the instrument was purchased by the Psychological Corporation and was renamed MicroCog (Powell et al.). In the new instrument, the number of items was reduced and response-time was incorporated in the scoring (Douglas Powell, personal

communication, November 4, 2003). The MicroCog (Powell et al.) was also normed as a general neuropsychological screen (Kane, 1998).

In summary, the MicroCog (Powell et al., 1993) is a computer-administered screening assessment that has the capacity to detect dementia and other neuropsychological impairment in an older population. However, its ability to detect cognitive decline at other ages, or to differentiate between dementia and other mental disorders has not been well established (Elwood, 2001). There is a need for continued research to better understand the limits and strengths of the MicroCog (Powell et al.). Having been developed for physicians, "Microcog shows virtually no ceiling effect and thus can detect cognitive deficits in well-educated, higher functioning individuals" (Elwood, 2001, p. 99). The MicroCog (Powell et al.) was designed as a screening device and was not intended to replace a full neuropsychological evaluation. However, as a screening instrument the MicroCog (Powell et al.) provides clinicians with a valid, reliable, and well normed assessment tool (Elwood; Green et al., 1994; Powell et al).

Demographic Form

The purpose of the questionnaire was to gather demographic and related information. Items include: (a) gender, (b) age, (c) ethnicity, (d) number of years the physician has been practicing medicine, post-residency, (e) specialty,

(f) current job status, (g) any known neuropsychological impairment, and (f) level of experience with computers (see Appendix C).

Procedure

Two sets of data were used in this study. The first set was obtained from the Center for Personalized Education for Physicians (CPEP), which has data on approximately 361 physicians who were given the MicroCog (Powell et al., 1993) as a part of their competency evaluation. No identifying information was attached to the MicroCog scores. However, demographic information including age, specialty, gender, handedness, and reason for referral were collected from individual files and placed in a database, along with nine index scores from the MicroCog (both reference scores and demographically adjusted scores; Powell et al.). These data were used in analyses, and compared with the second set of data that was collected specifically for this project.

The second set of data was collected by this researcher. Physicians for this study were recruited via three sampling strategies. The estimates of the number of participants recruited from the following sampling procedures were based on Powell and Whitla's (1994) response rates. Powell and Whitla used recruitment methods similar to this study to find physicians to take the ACS (the older version of the MicroCog). Of the 18,831 physicians invited to participate by Powell and Whitla, 12.4% (2,342) responded, 6.6% (1,246) made an appointment for testing, and 5% (908) completed the testing.

The following recruiting strategies were used in the order listed. The first method of recruiting participants for this study was snowball sampling, which means that physicians and other professionals (e.g., psychologists) known to this researcher were asked for help in supplying names of physicians who might be interested in this study. This method was also used in conjunction with the other two sampling strategies. Each physician who participated in the study was asked to mention the study to other physicians who might be interested in participating and were provided with three business cards listing the topic of the study and the name and phone number of the principal investigator (Appendix D). Four physicians responded to business cards given to them by their peers. Of these 4, 75% (3) scheduled a testing appointment and completed testing. None of the physicians tested through snowball sampling were known to this researcher.

For the second sampling strategy, permission was obtained from the University of Colorado Health Sciences Center (UCHSC) and Denver Health Medical Center (DHMC), and e-mails were sent to physicians informing them of the study and asking for volunteers (Appendix E). After determining that physicians 50-80 years of age were responding in fewer numbers, a second email was sent to the physicians at DHMC, focused on that age group specifically (Appendix F). The number of physicians who received an email through the UCHSC was unknown. However, 24 physicians responded to the

UCHSC email, and 50% (12) physicians made an appointment for testing and completed testing. At Denver Health 283 physicians were emailed with a 10.6% (30) response rate, and 8% (23) made an appointment and completed testing. To recruit physicians from DHMC and UCHSC a second review board approval was required. Please see Appendix A for Colorado Multiple Institutional Review Board approval.

Third, physicians who were paid consultants with CPEP were invited to participate in this study. These physicians were "clinical consultants," who assisted in competency evaluations by reviewing charts and conducting oral interviews with the physicians being evaluated (E. Korinek, personal communication, January 19, 2004). In their work with CPEP, the consultants were completely separate from the portion of the evaluation that involved the MicroCog (Powell et al., 1993). These physicians received a letter of invitation to participate, which included a description of the study (see Appendix G), a participation form (see Appendix H), a letter of support from CPEP (see Appendix I), and a pre-addressed response envelope. Of the 194 CPEP-affiliated physicians invited to participate, 43.8% (85) responded, 16.5% (32) made an appointment for testing, and 16% (31) completed the testing.

For this current study, a total of 143 physicians responded to an invitation to participate, 70 (49.0%) made an appointment, and 69 (48.3%) completed testing. The percent of physicians (49.0%) who made an appointment after responding to an invitation to the study is similar to the

percent found in Powell and Whitla's (1994) study where 53.3% made an appointment. Seventeen (11.9%) of the total physicians who responded to an invitation were eliminated due to exclusion criteria. Another 26 (18.2%) physicians responded to the invitation indicating that they were not interested in participating in the study, 15 (10.5%) did not respond after initially expressing interest, 1 (0.7 %) physician did not want to take a cognitive-based test, and for 14 (9.8%) either did not provide a reason or expressed other reasons. Some of the reasons provided by physicians who responded to an invitation but did not schedule an appointment included "only curious about the study and did not want to participate," "could not find time in their schedule to take the test," and "personal issues such as illness or family concerns."

Administration of the MicroCog (Powell et al., 1993) took place at several locations in Denver, Boulder, Greeley, and Colorado Springs. Local libraries, universities, and private offices were used as testing sites.

Physicians received information about this study through the sampling strategies discussed in the prior section. The recruitment advertisements included key pieces of information concerning the study, length of testing, and confidentiality information so that the participant began the informed consent process upon reading the advertisements (see Appendices E, F, and G). Once physicians replied to the correspondence, they were contacted and provided with more information concerning the study, informed of exclusion criteria, and if physicians continued to express interest and did not meet any exclusion

criteria, an appointment was scheduled for testing at a site most convenient to them.

This researcher met the participant outside the testing site and escorted him or her to the testing room. During the study, it was discovered that one participant had not been born in the United States. Since he or she had already taken the test, the participant was given the compensation and the incident was reported to both institutional review boards. To address any future difficulties, this researcher started re-qualifying participants when they arrived for their appointment.

The formal assessment of informed consent occurred when the examiner instructed the participant to read the informed consent form (see Appendices J, K, and L). Once the participant read the consent form, the examiner verbally reviewed the critical items on the consent form (confidentiality, how the data would be used, length of session, and ability to withdraw from study at any time) and engaged the participant in a discussion to assess comprehension. During this discussion the participant was asked to explain the purpose of the study in his or her own words. At this point the examiner gave the participant time to consider whether he or she wanted to continue with the study. If the participant decided to participate in the study, the examiner asked him or her to complete the Authorization B: Enrollment into Research HIPPA form (see Appendix B, for participants from COMIRB sites only), the demographic

questionnaire (see Appendix C), and the Physician Information Sheet (see Appendix M). A copy of the informed consent form and the Authorization B form (for participants from COMIRB sites only) were provided for the participant upon the conclusion of the session. All physicians chose to participate in and completed the study.

The participants were reminded that they would receive feedback on their performance via postal mail if their performance was in the expected range (see Appendix N) or by telephone if their performance was outside the expected range (see Appendix O for phone script). Names of all participants went on a list that allowed this researcher to associate names with participant test numbers. The list of names of physicians with their associated test numbers was kept in a locked filing cabinet in an undisclosed location during the study. Once individual feedback was provided or 30 days after the testing session (whichever came first), the participant's name was removed from the list matching his or her identification number with his or her name. Upon completion of data collection and feedback, all identifying data were destroyed.

Once the initial paperwork described above was completed, the examiner provided an explanation of the testing procedures and familiarized the examinee with the appropriate keys and procedures as outlined in the MicroCog manual (Powell et al., 1993). The MicroCog was administered on a laptop computer with an external keyboard and speakers attached. The physician sat at a desk or table with the computer in front of him or her to complete the test.

The examiner remained in the testing room for the entire administration to answer any questions the participant had, as recommended by Powell et al. The MicroCog (Powell at al.) took 45 to 60 minutes to complete.

After the physician completed the assessment, the examiner saved the test results on the computer under the appropriate participant number. Identification numbers were used to protect the confidentiality of the participants. The physician was given $50 in appreciation for his or her time or a donation of $50 was made to a charity. A boxed meal was provided after the testing session. The examiner then thanked the participant, provided copies of the informed consent and HIPPA Authorization B (for COMIRB participants only) forms, and ended the session. Once the study was completed, group results were sent to all requesting physicians.

For physicians who performed within the expected range, this researcher mailed a form letter within two weeks of the testing session. Please see Appendix N for the form letter. None of the physicians in the control group tested outside the expected range. However, there was a procedure delineated had a participant's performance been outside the expected range. If a physician's scores had indicated possible impairment, this researcher would have contacted the physician by telephone and provided the information outlined in Appendix O along with the recommendation that they contact the Colorado Physicians Health Program for assistance in obtaining a confidential neuropsychological evaluation. The Colorado Physicians Health Program

(CPHP) is an agency designed to provide confidential support to physicians who are having difficulties in their practice. Therefore, this agency would have been ideal for providing support to physicians whose test results exhibited deficiencies in function. Through the CPHP, participants from this study could have obtained confidential neuropsychological testing, rehabilitation, and counseling.

Approximately 15 (22%) physicians asked for more specific feedback. This researcher set up a time to provide individual feedback and reviewed the results of the testing with the participants.

This investigator conducted the testing. Before testing, the examiner received training in administering the assessment, in maintaining confidentiality, in assessing comprehension of informed consent, and took the MicroCog (Powell et al., 1993) assessment herself. Dr. McRae and Dr. Thompson supervised this investigator.

Analysis of Data by Research Question

In this section, the research questions are presented and the analyses that were conducted to address each question are described. There were 267 MicroCog (Powell et al., 1993) assessments available from the physician group who completed competency evaluations. Sixty-nine physicians were tested in the control group. One participant's data was eliminated because he or she did not meet the exclusion criteria. Therefore, 68 physicians were included in the control group. The alpha level was set at .05 for the following hypotheses.

Hypothesis #1: As measured by the MicroCog (Powell et al., 1993), the physician group referred for competency evaluations will score significantly lower than the control physician group (those who have not been referred for competency evaluations) on measures of speed, accuracy, and cognitive proficiency.

Three two-tailed t-tests were employed to address this question with physician group as the independent variable and speed, accuracy, and proficiency scores as the dependent variables. For a power level of .80, alpha = .05, a sample size of 26 was needed, assuming a large effect size, and 64 for a medium effect size. Prior to data collection harmonic n' was calculated at 77, ES = 1.86 standard units (calculated effect size from Williams, Williams, & Norcross, 2002), and alpha = .05, resulted in power > .99; for ES = .5 (medium effect size) power = .92 (Cohen, 1988, 1992). Age between the two samples was not significantly different; therefore, age was not used as a covriate.

Hypothesis #2: The proportion of physicians with cognitive impairment in the group of physicians referred for competency evaluations will be significantly greater than the proportion of physicians with cognitive impairment in the control group of physicians.

Neuropsychological impairment was determined by a General Cognitive Proficiency Index score of greater than one and one half standard deviations

below the mean (standard cutoff score = 77.5), or any two index scores that were one and a half standard deviations below the mean. A Chi-Square test was employed to determine if there was a significant difference between these two groups. Power calculation prior to data collection included alpha = .05, ES = .2 (calculated effect size from expected values), and N = 337, resulting in power > .88 (Cohen, 1988, 1992).

Hypothesis #3: As measured by the MicroCog (Powell et al., 1993), the magnitude of difference on the mean scores of the General Cognitive Proficiency Index, between the two physician groups will increase as physician age increases.

A two-way analysis of variance was employed to address this hypothesis, with age group (4 levels) and physician group membership (2 levels) as the independent variables, and the General Cognitive Proficiency score was the dependent variable. Prior to data collection power calculations were as follows; for a power level of .80, alpha = .05, a sample size of 13 per cell, or a total of 104 would be needed, assuming ES = .40 (a large effect size), and 240 for ES = .25 (a medium effect size; Cohen, 1988, 1992, Faul & Erdfelder, 1992). Schaie and Willis (1993) conducted an analysis of variance for a similar question and calculated large effect sizes.

Hypothesis #4: The control physician group will score significantly higher than the age-corrected norms at the highest educational level for the MicroCog (provided in the manual; Powell et al., 1993) on measures of speed, accuracy, and general cognitive proficiency.

Three one-sample t-tests were used to analyze this data. An analysis of power prior to data collection concluded that with n = 45, d = .99 (calculated based on expected difference for General Cognitive Proficiency Index score) standard units, and alpha = .05, would result with power > .99; and for d = .70, power = .90 (Cohen, 1988, 1992).

CHAPTER IV

RESULTS

Introduction

This chapter includes an initial comparison of the mean age for the two physician groups to determine if age would need to be covaried in the following analyses. Next is a presentation of specific data analyses and results for each hypothesis put forth in this study. Also, there is a section where additional analyses and results are presented. The additional analyses include gender differences, a comparison of the control group of physicians and the group referred for competency evaluation with all impaired physicians (as defined by hypothesis three of this study) removed from the data, and comparisons of five index scores between the two physician groups. Finally, to explore the question of using age-corrected norms for determining physician competency mentioned by Turnbull et al. (2000), the results for age groups of the two physician groups using a normative reference group without age- or education-correction on the proficiency score are presented.

Data Analyses

Prior to analysis all variables were reviewed for data entry errors and missing data. Missing scores for one domain index were found for one

participant. The index score was used in two analyses; thus, this case was deleted from those analyses. Entries were determined to be within an acceptable accuracy range, by reviewing extreme values for each variable used. All outliers were viewed individually and checked for accuracy. Unless otherwise noted, the scores used for all analyses were age- and education (higher than high school)-corrected. An alpha level of .05 was used for all statistical tests. Statistics were calculated using SPSS Graduate Pack 13.0 for Windows (2004). Pallant's (2001) book, SPSS Survival Manual, was used for information concerning data analyses with SPSS, writing results, and using graphs.

Age Difference Between Two Physician Groups

The statistical assumptions of independence and normality were met. The data contained no outliers, as identified by the SPSS Boxplot. Normality was assessed using skewness and kurtosis measures. However, homogeneity of variance was not met. Greater variance occurred in the control group which had a smaller sample size. This occurrence violated one of the assumptions for using a t-test, creating the possibility that the actual alpha may exceed the nominal alpha. Thus, a more stringent alpha level of .01 was used for this analysis. In consequence, the results reported were from the "equal variances not assumed" category.

To determine if age of the competency evaluation group of physicians (n = 267, M = 51.45, SD = 9.11) was different from the control group of physicians (n = 68, M = 49.40, SD = 12.45), a two-tailed, independent-samples t-test was conducted. There was no statistically significant difference between the mean age of the two physician groups, $t(86.13) = -1.27$, $p = .206$. The effect size for this analysis was very small (eta squared = .005; Cohen, 1988; Pallant, 2001). Since there was no statistically significant difference between the mean age of the two physician groups, age was not covaried in the following analyses.

Hypothesis 1

As measured by the MicroCog (Powell et al., 1993), the physician group referred for competency evaluations will score significantly lower than the control physician group (those who have not been referred for competency evaluations) on measures of speed, accuracy, and cognitive proficiency.

Prior to each analysis, statistical assumptions of normality, homogeneity of variance, and independence were checked. Each variable used in this section demonstrated independence. Skewness measures suggested that outliers were impacting normality, thus outliers were removed. Outliers were identified with the SPSS Boxplot. The number of outliers removed for each variable were as follows: speed - competency group 4 and control group 1; accuracy - competency group 5 and control group 1; and proficiency - competency group 2

and control group 0. After eliminating outliers, each variable in these analyses was determined to be within an acceptable range of normality, using skewness and kurtosis measures. All t-tests were calculated both with and without identified outliers, and both sets of results were statistically significant. The statistics reported here were from analyses with the outliers eliminated.

Homogeneity of variance was assessed using Levene's Test for Equality of Variances. For the variables that were determined not to have equal variance, the variance was greater in the larger group. According to Glass and Hopkins (1996), if the greater variance is in the larger group, the "true probability of a type-I error is always less than the nominal probability"(p. 293). Therefore, in this circumstance the violation of the homogeneity of variance was of no concern.

An examination of whether there were differences between a control group of physicians and physicians referred for the competency evaluation revealed that there were significant differences on measures of speed, accuracy, and cognitive proficiency (see Table 4 for means, standard deviations, and mean differences for speed, accuracy, and proficiency comparisons).

Speed

A two-tailed, independent-samples t-test demonstrated that there was a significant difference in the mean Processing Speed Index score between the control group of physicians and the competency evaluation physicians, $t(161) =$

7.5, p < .001. The magnitude of the difference in the means, in terms of the effect size, was large (eta squared = .146; Cohen, 1988; Pallant, 2001).

Accuracy

A two-tailed, independent-samples t-test showed that the mean Processing Accuracy Index score of the control group of physicians' was significantly higher than the competency group's, t(327) = 4.275, p < .001. The magnitude of the difference in the means was moderate (eta squared = .053; Cohen, 1988; Pallant, 2001).

Cognitive Proficiency

Using a two-tailed, independent-samples t-test, a significant difference was found in the mean General Cognitive Proficiency score between the control group of physicians and the competency evaluation physicians, t(135) = 10.3, p < .001. The magnitude of the difference between the means was very large (eta squared = .244; Cohen, 1988; Pallant, 2001).

On all three t-tests the control group scored significantly higher than the competency evaluation group. The largest difference between groups was on the cognitive proficiency score, with a mean difference of 13.79. With the MicroCog's (Powell et al., 1993) standard score of 100 and standard deviation of 15, a mean difference of 14 indicates that there was a difference of almost one standard deviation between these two groups.

Table 4 Means, SDs, and Mean Differences Between Two Physicians Groups

	Control		Competency		
	M	SD	M	SD	Mean Difference
Speed	109.42	10.36	97.26	16.42	12.16
Accuracy	105.28	9.22	99.65	9.729	5.64
Proficiency	109.62	9.126	95.83	12.58	13.79

Hypothesis 2

The proportion of physicians with cognitive impairment in the group of physicians referred for competency evaluations will be significantly greater than the proportion of physicians with cognitive impairment in the control group of physicians.

Statistical assumptions for this analysis were met, in that, none of the expected values were less than 5. Neuropsychological impairment was determined by a General Cognitive Proficiency score of more than one and one half standard deviations below the mean (standard cutoff score = 77.5), or any two index scores that were one and one half standard deviations below the

mean. The observed values were determined by the number of individuals with and without neuropsychological impairment in each physician group. The observed values were entered into the Crosstabs analysis from SPSS 13.0, and expected values were calculated from this program.

A 2x2 Chi-Square analysis demonstrated that there was a significant difference between the proportion of physicians with cognitive impairment in the two physician groups, Pearson Chi-Square (335) = 10.92, $\underline{p}$ = .001. Please see Table 5 for expected and observed values. In the competency evaluation group, 14.6% of the physicians were impaired, compared to the control group where 0% were impaired. The effect size for this test was small ($\underline{w}$ = .163; Faul & Erdfelder, 1992).

Hypothesis 3

As measured by the MicroCog (Powell et al., 1993), the magnitude of difference on the mean scores of the General Cognitive Proficiency Index, between the two physician groups will increase as physician age increases.

The statistical assumptions of normality, homogeneity of variance, and independence were met for this analysis. While there were outliers identified by SPSS Boxplot, their influence on their respective means was judged to be minimal, because all variables passed normality tests with outliers included. Therefore, outliers remained in the analysis.

Table 5 Expected and Observed Values in Two Physician Groups for Impaired Physicians

	Impaired	Unimpaired	Total
Control			
Expected	7.7	60.3	68
Observed	0	68	
Competency			
Expected	30.3	236.7	267
Observed	38	229	
Total			
Expected	38	297	335

A two by four between-groups analysis of variance was conducted to explore the impact of physician group and physician age group on the proficiency score. Participants were divided into four age groups (30 - 39; 40 - 49; 50 - 59; 60 and above). Physicians between the age of 30 and 80 participated in this research project. The ideal would have been to have age groups through the 8th decade. However, there were too few physicians in the oldest categories to form a group. Thus, the 60s, 70s, and 80s were collapsed

into a single category. Please see Table 6 for cell size, cell mean, and cell standard deviation.

Results indicated there was a statistically significant main effect for physician group, $F(1, 327) = 64.14$, $p < .001$, with a large effect size (partial eta squared = .164). There was no statistically significant effect for age group, $F(3,327) = .327$, $p = .806$ Also, the interaction effect was not significant, $F(3, 327) = 1.027$, $p = .381$. With the small effect size (partial eta squared = .009), a larger sample (N = 1,096; Cohen, 1988) would be needed to find an actual interaction. The mean differences between the control group and the group referred for competency evaluations are shown in Table 7.

Table 6 Descriptive Statistics for Analysis of Variance

Group	Age Group	N	Mean	SD
Control				
	30	14	110.21	10.05
	40	24	108.33	8.90
	50	17	110.00	8.97
	60 +	13	110.85	9.53
Competency				
	30	25	98.60	13.19
	40	89	98.09	12.98
	50	93	94.18	11.57
	60 +	57	93.86	12.83

Table 7 Proficiency Score Mean Differences by Age Group, Between Control and Competency Group Physicians.

Age Group	Mean Difference
30	11.61
40	10.24
50	15.82
60 +	16.99

Hypothesis 4

The control physician group will score significantly higher than the age-corrected norms at the highest educational level for the MicroCog (provided in the manual; Powell et al., 1993) on measures of speed, accuracy, and general cognitive proficiency.

Prior to each analysis the statistical assumptions of normality were assessed using skewness and kurtosis. After eliminating one outlier, which was identified by SPSS Boxplot, for the accuracy variable, normality was achieved for all variables in this section. The accuracy t-test was calculated both with and without the identified outlier; both results were significant. The statistics reported from the accuracy data were with the outlier eliminated. For the speed and proficiency data no outliers were found.

Three one-sample t-tests showed that the physician control group performed significantly higher than the norm group (age-corrected and education higher than high school) on measures of speed, $t(67) = 6.696$, $p < .001$; accuracy, $t(66) = 4.692$, $p < .001$; and proficiency, $t(67) = 8.69$, $p < .001$. The respective effect sizes were moderate ($d = .596$), between moderate and small ($d = .396$), and moderate ($d = .641$; Cohen, 1988). Please see Table 8 for means, standard deviations, and mean differences. The greatest mean difference, similar to the comparison to the competency group, was on the cognitive proficiency score.

Table 8 Control and Norm Sample Means, SD, and Mean Differences on Speed, Accuracy, and Proficiency Index Scores.

	Control		Norm Sample		
	M	SD	M	SD	Mean Difference
Speed	108.94	11.01	100	15	8.94
Accuracy	105.94	10.64	100	15	5.94
Proficiency	109.62	9.126	100	15	9.62

Additional Findings

Computer Experience Related to the Global Proficiency Index Score

Computer experience was a categorical variable with four levels (Very, Moderate, Some, and No). No participant selected the "No Experience" category. For this analysis, the "Moderate" and "Some" level of computer experience were collapsed into one category, "Less Experienced", due to small n's. Thus, the two levels of computer experience for this analysis were "Very Experienced" and "Less Experienced".

The statistical assumptions of independence, normality, and homogeneity of variance were met for this analysis. The one identified outlier was not removed from the data. It was judged to have minimal impact on the mean, and normality was achieved with it included in the data. All analyses were calculated with the outlier included in the data.

A two-tailed, independent-samples t-test was conducted to compare mean cognitive proficiency scores between control sample physicians who reported that they were very experienced with computers and those who reported less experience with computers. There was no significant difference between the two groups [$t(66) = .913$, $p = .364$]. Please see Table 9 for means, standard deviations, mean differences, and cell sizes. The degree of the differences in the means was small (eta squared = .012; Cohen, 1988; Pallant, 2001).

Table 9 Results of the Comparison of Control Group Physicians with Different Experience on Computers

Group	N	M	SD	M Differences
Very Experienced	48	110.27	9.06	2.21
Less Experienced	20	108.05	9.32	2.21

Group Differences with Cognitively Impaired Physicians Removed

For this analysis, all physicians who were identified as being cognitively impaired were removed from the sample, thus providing a comparison of unimpaired physicians from both groups. Neuropsychological impairment was determined by a General Cognitive Proficiency Index score of more than one and one half standard deviations below the mean (standard cutoff score = 77.5), or any two index scores that were one and one half standard deviations below the mean. The statistical assumptions of normality, homogeneity of variance, and independence were met for this analysis. One outlier was identified by SPSS Boxplot and was judged to impact normality; therefore it was removed from the data. The analysis was conducted with and

without outliers, both with the same significant results. The results reported here were with the outlier eliminated from the data.

A two-tailed, independent-samples t-test was employed to compare cognitive proficiency scores for a control group of physicians and a group physicians referred for competency evaluations (with impaired physicians removed). There was a significant difference [$\underline{t}(294) = 8.044$, $\underline{p} < .001$] between the control group ($\underline{n} = 68$, $\underline{M} = 109.62$, $\underline{SD} = 9.126$) and the unimpaired physicians referred for competency evaluations ($\underline{n} = 228$, $\underline{M} = 98.90$, $\underline{SD} = 9.790$). The mean difference between these two groups was 10.719 and the magnitude of this difference was large (eta squared = .180; Cohen, 1988; Pallant, 2001).

Gender Differences

Statistical assumptions of independence, normality, and homogeneity of variance were met for this analysis. Two outliers, as identified by SPSS Boxplot, were removed, one from each of the male and female categories of the competency evaluation group, because their influence on their respective means was judged to be substantial, impacting normality. Statistics were calculated with and without outliers, resulting in the same significant or non-significant results. The reported statistics were with the outliers eliminated.

A two-tailed, independence-samples t-test was used to determine if there was a difference between control group females ($\underline{n} = 27$, $\underline{M} = 106.52$, $\underline{SD}$

= 8.437) and males (n = 41, M = 111.66, SD = 9.082) on a global cognitive proficiency score. A significant difference [t(66) = 2.198, p = .O.022] was found between the two groups, with the males scoring higher than the females. The mean difference between the two groups was 5.14, and the magnitude of difference between these two groups was moderate (eta squared = .077; Cohen, 1988; Pallant, 2001).

A different result was found when this same comparison was made between genders in the competency evaluation group. There was no significant difference [t(263) = -1.555, p = 121] between males (n = 222, M = 95.30, SD = 12.371) and females (n = 43, M = 98.44, SD = 10.804) on a global proficiency score. The mean difference between the two groups was 3.15, with the degree of difference being small (eta squared = .009; Cohen, 1988; Pallant, 2001).

Group Differences on Five Neuropsychological Domains

Prior to each analysis, statistical assumptions of normality, homogeneity of variance, and independence were checked. Each variable used in this section demonstrated independence. Homogeneity of variance was assessed using the Levene's Test for Equality of Variance. For the variables that were determined not to have equal variance, the variance was greater in the larger group, in which case the actual alpha may be less than the nominal alpha. Normality was assessed using skewness and kurtosis measures. Due to negative skewness, it was determined that outliers identified by SPSS Boxplot were

impacting normality, and thus, they were removed. All t-tests were calculated both with and without identified outliers; both results were significant. The statistics reported were from data with the outliers eliminated.

After eliminating outliers, the Attention/Mental Control, Reasoning/Calculation, and Memory variables were assessed as being within an acceptable range of normality. To address the concern of inflation of Type I error, alpha was reduced to .01 for these three t-tests.

For the remaining two variables, Spatial and Reaction Time, even with outliers removed, the data for the competency evaluation group remained negatively skewed. To address this concern, Tabachnick and Fidell (1996) reported that in large samples ($n > 200$) the impact of skewness diminishes. Also, for these two variables a smaller alpha of .001 was used.

Four of the two-tailed, independent-samples t-tests showed that the physician control group performed significantly higher than the competency evaluation group [Attention/Mental Control, $t(125) = 9.162$, $p < .001$; Reasoning/Calculation, $t(332) = 3.792$, $p < .001$; Memory, $t(130) = 5.787$, $p < .001$, and Spatial Abilities, $t(127.7) = 7.479$, $p < .001$]. The respective effect sizes were very large (eta squared = .20), moderate (eta squared = .04), between moderate and large (eta squared = .09), and large (eta squared = .14; Cohen, 1988). The greatest mean difference was on the Attention/Mental Control Index.

Results of the two-tailed independent-samples t-test for Reaction Time indicated that there was no significant difference between the control physicians and the competency evaluation physicians, $t(142.55) = .780$, $p = .436$. The effect size for this analysis was very small (eta squared = .002). Please see Table 10 for means, standard deviations, and mean differences.

Table 10 Means, SDs, and Mean Differences of Two Physician Groups on Five Domains.

	Control		Competency		
Index	M	SD	M	SD	Mean Difference
Attention/Mental Control	110.07	9.09	98.15	11.25	11.92
Reasoning/Calculation	106.87	12.27	99.88	13.86	6.98
Memory	110.41	10.54	101.58	13.58	8.93
Spatial Abilities	108.79	8.26	99.90	10.45	8.89
Reaction Time	105.93	7.85	105.02	10.95	.677

Physician Proficiency Scores Without Age- or Education-Correction

Unlike the scores used for all previous analyses, this analysis incorporated proficiency scores for both physician groups that were neither age- nor education-corrected. The normative reference group for this analysis was made up of adults between the ages of 18 and 34, with no educational level information used (Powell et al., 1993). Table 11 provides the number in cell, means, and standard deviations for the two physician groups. The mean differences between the control group and the group referred for competency evaluations were 30s - 13.10, 40s - 10.30, 50s - 14.41, and 60s - 9.21. Also, Figure 3 illustrates the mean scores without age-correction.

Table 11 Physician Proficiency Scores Normed on a Reference Group with No Age Correction

Group	Age Group	N	Mean	SD
Control				
	30	14	114.14	10.48
	40	24	108.54	7.52
	50	17	105.00	7.75
	60	13	92.23	8.98
Competency				
	30	25	101.04	12.62
	40	89	98.24	11.66
	50	93	90.59	10.61
	60	57	83.02	11.45

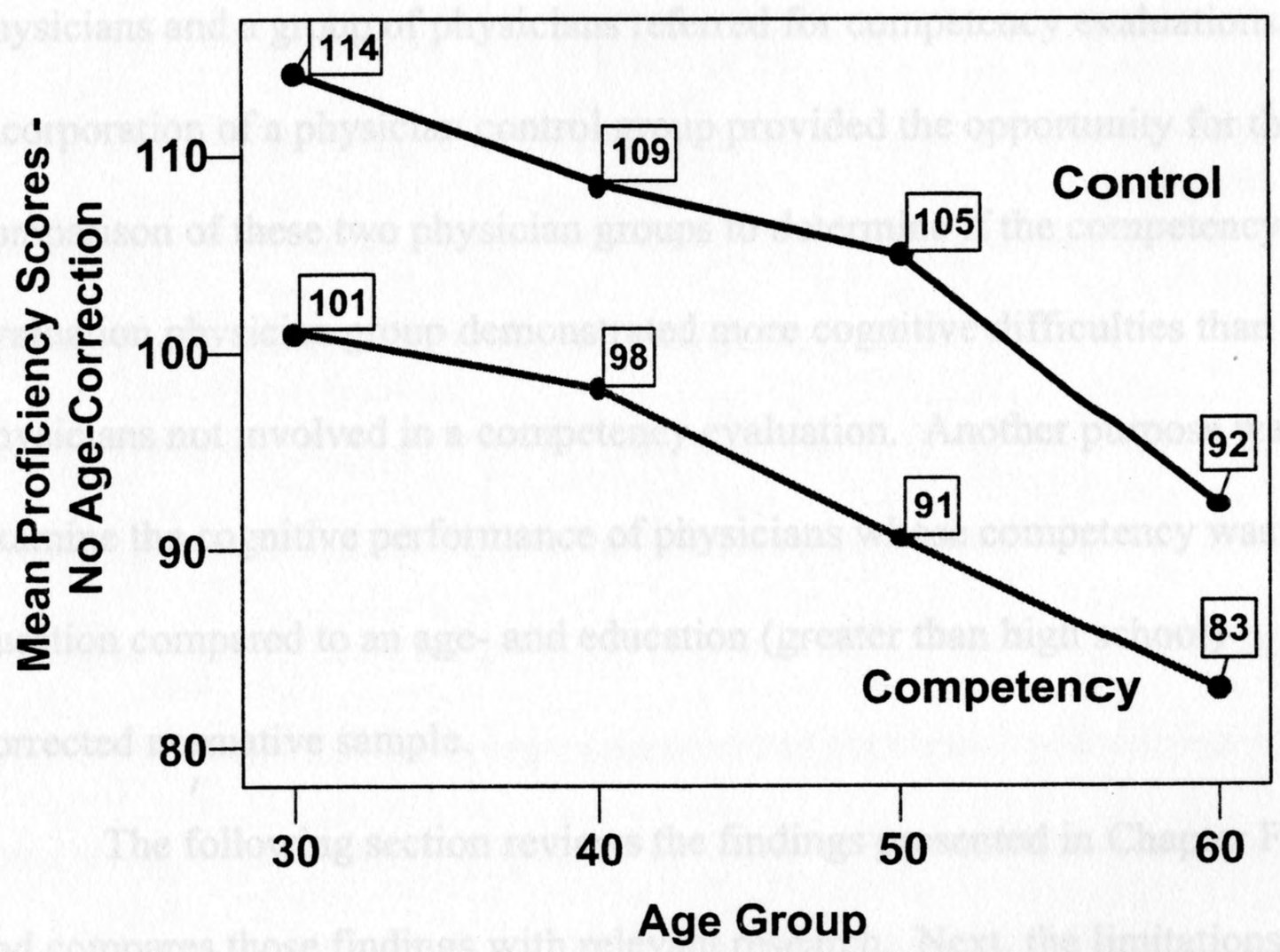


Figure 3 Proficiency Means for Physician Age Groups No Age-Correction

CHAPTER V

DISCUSSION

Introduction

The purpose of this study was to learn more about the neuropsychological characteristics of two physician groups; a control group of physicians and a group of physicians referred for competency evaluations. The incorporation of a physician control group provided the opportunity for the comparison of these two physician groups to determine if the competency evaluation physician group demonstrated more cognitive difficulties than physicians not involved in a competency evaluation. Another purpose was to examine the cognitive performance of physicians whose competency was not in question compared to an age- and education (greater than high school)-corrected normative sample.

The following section reviews the findings presented in Chapter Four and compares those findings with relevant research. Next, the limitations of this research project are presented. Finally, potential implications for the medical field are explored, and future directions for research are discussed.

It is imperative to understand that this discussion is about group differences and not about individuals. There were several individuals from the

competency evaluation group who performed exceptionally well on the MicroCog (Powell et al., 1993). Therefore, it would be incorrect and harmful to assume that any physician referred for a competency evaluation would perform the same as others referred for competency evaluations. Reasons for referral vary widely and may or may not involve a neuropsychological component.

Discussion of Results

Hypothesis 1

The findings for the first hypothesis for this study indicated that the control group of physicians performed better than the physicians referred for competency evaluations on measures of information processing speed, information processing accuracy, and cognitive proficiency.

While all of the differences were significant, the largest effect size was found in the comparison of proficiency scores (eta squared = .244). The proficiency score from the MicroCog (Powell et al., 1993) is a global cognitive functioning score that incorporates measures of both speed and accuracy, giving processing accuracy more weight. The difference between the two groups was almost one standard deviation. This finding is similar to the only other study known to this author that used a control group (n = 9). Williams et al. (2002) reported differences between two similar physician groups on two subtests from the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1987; vocabulary and picture arrangement) and the Wide Range Achievement Test-Third Edition

(Wilkinson, 1993). The differences appeared to be about one standard deviation.

On the speed of processing measure, the difference between the two physician groups in this study also had a large effect size (eta squared = .146). There were no other studies known to this author available for comparison of this result. However, a mean difference of 12 with a population standard deviation of 15 is a large difference. Processing speed is one of the measures that is particularly sensitive to neuropsychological impairment (Gronwall, 1987; van Zomeren & Brouwer, 1992). It appears that as a group, physicians referred for competency evaluations tend to process information more slowly than a control group of physicians.

For the accuracy score, the mean difference of 5.6 between the two groups had the smallest effect size of the three measures tested. While accuracy can be adversely impacted by cognitive impairment, in this sample of two physician groups it may not have been as sensitive to cognitive impairment as processing speed.

It is interesting to note that some of the control group physicians expressed that they thought that there would be no difference in the cognitive constellation of the control group and the physicians referred for competency evaluations.

Hypothesis 2

A significant difference was found in the proportion of physicians with scores that suggested cognitive impairment, between the control group of physicians and the group of physicians referred for competency evaluations. Cognitive impairment was defined as performing one and a half standard deviations below the mean, or lower, on a global cognitive proficiency score, or on any two index scores. Using this definition, none of the physicians in the control group demonstrated cognitive impairment. However, 14.6% (n = 38) of the physicians in the competency evaluation group had scores that suggested cognitive impairment. This finding is less than what has been found in similar research. Turnbull et al. (2000) found that of the 27 physicians evaluated in their study, 26% demonstrated cognitive impairment. The finding in this current research could be different due to the different pathways in which physicians were referred to these programs. For the Turnbull et al. study, physicians only received the neuropsychological tests after being identified and provided with an opportunity to improve their performance. This opportunity may have provided a filter by which impaired physicians without cognitive difficulties could have avoided the neuropsychological evaluation. Thus, fewer physicians without cognitive difficulties actually reached the evaluation stage. Also, Turnbull et al. used traditional neuropsychological tests. The current study used an hour long neuropsychological screen. It was not designed to

replace a full battery of neuropsychological tests. Such a battery might provide greater opportunity to find cognitive impairment. More information is needed on how this particular neuropsychological screen compares to traditional batteries in identifying impaired examinees. Another possible reason for the difference is that this study excluded physicians who were foreign born, foreign educated, and whose primary language was not English, because of their potential for performing poorly on English-based cognitive tests (Jacob et al., 1997). Turnbull et al. did not mention such exclusion criteria.

Another study that did not mention exclusion of specific physician groups was Thompson's research (2003). Like the current study, Thompson used data from CPEP for her analysis. While there was an overlap of participants between these two studies, the current study incorporated physicians that were tested since Thompson's work. She found that 25% of the physicians scored low enough to warrant concern about their cognitive abilities. The criteria for determining cognitive impairment tend to involve some subjective interpretation and were not operationalized or clearly specified in Thompson's and Turnbull et al.'s studies. Therefore, it is difficult to compare how neuropsychological impairment was determined. The findings of Turnbull et al. and Thompson were very similar.

A more recent study by Perry and Crean (2005) defined cognitive impairment as scoring one standard deviation below the mean on

neuropsychological measures. Since this current study used one and one half standard deviations as a cutoff point for establishing cognitive impairment, the different cutoff scores make it difficult to compare results. Perry and Crean found that 51% of the physicians in their competency evaluation group performed in the cognitively impaired range on four or more neuropsychological instruments. While this seems like a high number compared to the other studies, the authors included individual subtest scores from tests when assessing the number of cognitively impaired scores. This study only used index scores that were a composite of individual test scores. By the very number of scores available when individual subtest scores are used, the number of scores suggesting cognitive impairment using composite index scores would be expected to be less. As noted by the authors (Perry and Crean), their research project did not have a control group with which to compare results. Thus, it is difficult to know how many test scores one standard deviation below the mean would be expected for a control group of physicians.

Certainly, a detailed comparison of other studies to this current study is difficult due to differing instruments, cutoff scores, and methods for determining cognitive impairment. However, all of these studies clearly indicate that there is a high level of neuropsychological impairment in a group of physicians referred for competency evaluations. With a control group for comparison this study extends that conclusion to say that there is more

neuropsychological impairment, as defined by this study, in a group of physicians referred for competency evaluations than in a group of physicians not involved in a competency evaluation.

Hypothesis 3

There was no significant statistical support for the hypothesis which stated that there would be an increase in the magnitude of difference of mean proficiency scores between the two physician groups with increases in age.

Certainly, many older physicians in the control group tested extremely well compared to their own age group from the MicroCog norms (Powell et al., 1993). Even when compared to the reference group (between the ages of 18 and 34), many older physicians' cognitive functions were well within the average to above average range. This result is similar to Powell and Whitla (1994) who noted that many older physicians' neuropsychological functioning was comparable to younger physician groups.

Hypothesis 4

Due to extensive education and high IQ (Matarazzo & Golden, 1972), it would be logical to assume that physicians' performance on neurocognitive tests would be higher than a normative sample of the general population. Hypothesis four explored this possibility by suggesting that the performance of the physician control group would be significantly better than the current age-corrected, greater than high school normative scores supplied by the MicroCog

(Powell et al., 1993). As with many neuropsychological tests, there is no MicroCog normative group specifically for physicians, and the education-correction may not accurately reflect physicians' level of education, intelligence, and experience. According to the results of this study, the control group of physicians performed significantly higher than an age- and education (greater than high school)-corrected normative group on measures of speed, accuracy, and proficiency. Perry and Crean (2005) mentioned their concern that on many neuropsychological tests physicians are being compared to normative groups that do not provide a fair representation of their intellectual and neuropsychological abilities.

Research on physicians' neuropsychological functioning is very limited. Powell and Whitla (1994) used an early version of the MicroCog (Powell et al., 1993) to collect information about the cognitive characteristics of physicians. In their comparison of physicians, practicing and retired, to a group of "normal" non-physicians, the physicians scored higher on a global accuracy score than the non-physicians in all age groups. This is similar to the results found in this current study. Powell and Whitla also noted that some of the differences in the groups could be attributed to the difference in the level of education and IQ.

Computer Experience Related to the Global Proficiency Index Score

The question of how using a computerized administration of cognitive assessment impacts test performance continues to be a concern. For this study,

using the control group, a comparison of physicians who reported being "Very Experienced" with computers and physicians who reported being "Less Experienced," found no significant difference on a global cognitive proficiency score. Computer experience was a categorical variable with four levels (Very, Moderate, Some, and No) that was collapsed into two levels of computer experience, "Very Experienced" and "Less Experienced".

The lack of difference could be attributed to the fact that all of the physicians tested reported having some experience with the computer. Since most physicians either worked in a setting where they used computers or at least had a home personal computer, taking a test on a computer might have been less of a concern for this sample than it was with earlier studies. Again, if physicians with no computer experience were compared to physicians with some computer experience, a difference might have been found.

Group Differences with Cognitively Impaired Physicians Removed

Physicians who met this current study's definition of cognitive impairment were removed from the sample. A comparison was then made between the control group of physicians and the remaining physicians referred for competency evaluations. This was done to determine whether the cognitively impaired physicians in the competency evaluation group accounted for the difference between the two groups. Analyses demonstrated that the control group still performed significantly better than the competency

evaluation group, even after removing the cognitively impaired physicians from the sample. The mean difference between the two groups was approximately two-thirds of one standard deviation at 10.71, compared to the mean difference of 13.79 with the cognitively impaired physicians included in the competency evaluation group. The competency evaluation group (without impaired physicians) still performed below the 50th percentile with a standard score of 98.90. This result indicated that the cognitive constellation of physicians who were referred for competency evaluations was distinctly different from a control group of physicians, as suggested by Perry and Crean (2005) and Thompson (2003).

Gender Differences

In their research with physicians, without separating age groups, Powell and Whitla (1994) did not find a significant difference between males and females on a global accuracy score on the MicroCog (Powell et al., 1993). The results from this current study were inconsistent. A significant difference was found between males and females in the control group. However, there was no significant difference between the two genders in the competency evaluation group. Due to the lower number of females in both groups and lack of random sampling, either result could be considered preliminary, but not particularly noteworthy.

Group Differences on Five Neuropsychological Domains

An exploration of the performances of the two physicians groups on separate neuropsychological domains revealed that four scores appeared to follow the same pattern found in the speed of processing, accuracy, and proficiency scores. The domains of Attention/Mental Control and Spatial Abilities had differences with large effect sizes, with a difference of approximately two-thirds of one standard deviation between the control physicians and the competency physicians. Reasoning/Calculation (language based and non-language based abstract reasoning) and Memory demonstrated differences with moderate to large effect sizes (ranging from one-half to two-thirds of one standard deviation) between the two groups.

While unable to compare the competency evaluation referred physician group to a control group in their study, it appears that Perry and Crean (2005) found deficits in their sample in areas similar to this study. Using traditional neuropsychological tests, these authors noted deficits in complex reasoning, memory, attention to detail, and spatial abilities.

For this current study there was only one neuropsychological domain, Reaction Time, in which there was no significant difference between the two groups. The lack of difference in a pure reaction time score is not surprising, in that the reaction time score was based on very simple tasks, pressing the enter key when a square was presented, or when a sound was heard. Simple reaction

time scores are not as discriminating as reaction time on more complex tasks, such as having to choose between two stimuli (Gronwall, 1987). Thus, greater differences between the two physician groups would be expected on more complex processing tasks, as seen on the Information Processing Speed Index.

Physician Proficiency Scores Without Age- or Education-Correction

All of the results discussed up to this point have been with the physicians' scores adjusted for age and education (greater than high school education). When assessing neuropsychological impairment, age corrections are used because cognitive functioning tends to decline with age. Therefore, adjustments to performance scores are made to correct for the decline in performance with increasing age. Turnbull et al. (2000) questioned this practice when assessing physician competency, suggesting that no matter the age, physicians need to demonstrate a certain level of cognitive ability to maintain a safe and competent practice. To address this concern, for this analysis only, both physician group scores were not adjusted for age.

The MicroCog (Powell et al., 1993) provides a normative reference group made of up 180 individuals between 18 and 34 years old, with no information about education levels. According to Powell et al., "The individuals in this age range generally achieved the highest scores in the sample" (p. 59).

The mean proficiency score (M = 83) for the group of competency evaluation physicians 60 years old and greater, was over one standard deviation below the mean, while their age- and education- corrected mean score (M = 93.86) was within the average range. This result highlights the discussion addressed by Turnbull et al. (2000), who questioned the use of age-corrected norms for physician competency evaluations. Perhaps using norms based on physicians' optimum functioning age group might provide a better standard when determining whether a physician can safely practice medicine. If older physicians are compared to their own age group, undetected cognitive deficits in their normal cognitive decline might adversely impact their ability to provide quality patient care.

Limitations of this Study

Caution is always wise when interpreting the findings of any study. This study has limitations that need to be considered when interpreting the results. While this study included a control sample as suggested by past researchers in this area (Perry & Crean, 2005; Thompson, 2003), the control sample may not have been representative of the physician population. This was not a ramdom sample of physicians. Participants were recruited from a city hospital, an academic setting, and from private practice in an effort to fairly represent the physician population. It is possible that the physicians who volunteered for the study were higher functioning individuals than the general

physician population. The physician control groups' scores appeared to be comparable with the scores of other physician control groups (Powell & Whitla, 1994; Williams et al., 2002). Other control groups were also volunteer, thus it is possible that all control groups were made up of higher functioning individuals or were biased in some way.

Another limitation to consider when interpreting findings from this study is that some physicians responded to an invitation to participate but never scheduled an appointment. It is possible that once these physicians understood that cognitive testing was involved, they eliminated themselves. Only one potential participant specifically stated that she or he did not want to take a cognitive test due to performance concerns. A total of 29 physicians self-eliminated due to over-packed schedules, physical illnesses, personal obligations, and for unknown reasons. Physicians that self-eliminated after indicating an interest could have been individuals with more cognitive difficulties than the group of physicians that did ultimately participate in the study. However, the variety of known reasons suggests no specific patterns indicating lower cognitive functioning in this group. Some of the physicians that were too busy were highly functioning individuals who were in high demand.

The control sample was collected from the Denver metropolitan area. Therefore, this group of physicians may not be representative of physicians

across the United States. Because many of the control group physicians grew up and were trained at locations across the United States, the sample could be more diverse than might be assumed with the small catchment area. However, caution should be used when generalizing this information to the entire population of physicians.

The instrument was a 60-minute neuropsychological screen. The very aspect that made collecting data from volunteers easier, could also contribute to limitations. This was not a full neuropsychological battery. A screen does not provide the depth of information that a full battery provides. Therefore, these results need to be interpreted with caution. An impaired score on the MicroCog (Powell et al., 1993) should not be viewed as a definitive result. Instead an impaired MicroCog score should be followed up with a full neuropsychological battery to determine and define possible cognitive impairment. However, the shorter time commitment required for the testing did facilitate recruitment of practicing physicians and the ability to test in a variety of settings.

Another limitation was introduced by the fact that the test was computer administered. While computers have developed to the point that they are valid administrators, there are still limitations. The MicroCog (Powell et al., 1993) is weak in presenting auditory stimuli (L. Thompson, personal communication, March 24, 2005) and measuring auditory-based neuropsychological domains. Most computer assessment programs are lacking in this area because the

technology to accurately measure verbal feedback has not developed to a practical level of use. Therefore, there is little information available on auditory learning, a modality used continually by physicians. However, a strength of using this computer generated assessment was that processing speed measures were available and incorporated into the scoring. Obtaining processing speed measures from human administered tests is difficult and imprecise at best. Measuring processing speed is highly advantageous because it is a neuropsychological function that is very sensitive to cognitive impairment (Gronwall, 1987; van Zomeren & Brouwer, 1992).

The difference in the testing situation of the two physician groups should also be taken into account when interpreting the results. Physicians in the competency evaluation group were not volunteers; they were required to have this assessment. The competency physicians were involved in an evaluation that could ultimately mean the end of their career and their livelihood. The anxiety and pressure created by that scenario could have negatively impacted their performance. On the other hand, the control group of physicians were actively recruited and were free to volunteer to take the test, with little fear that the results could mean the end of their practice. However, even in this group, several of the physicians expressed anxiety and concern over their performance. In fact, several volunteers indicated they wanted to be tested because of perceived changes in their own cognitive abilities. It is possible that

anxiety and concerns related to the testing scenario for the referral physicians contributed to some of the differences between the two groups. However, it is doubtful that the number and magnitude of statistically significant results in this study could be explained by these different testing scenarios.

Study Implications

The results of this study, along with several other studies (Madden, 1988; Perry & Crean, 2005; Thompson, 2003; Turnbull et al., 2000), suggest that, as a group, physicians referred for competency evaluations have more cognitive difficulties than physicians whose competency is not in question. If 14% - 26% of physicians referred for evaluation are cognitively impaired (Thompson 2003; Turnbull et al., 2000), then it is imperative that physician evaluation programs include neuropsychological evaluations. If the neuropsychological impairment comes from a progressive disease, then monitoring of abilities with a plan for early retirement could be put into place (Kapur, 1997). If there is a reversible problem, as suggested by Hanna et al. (2000) and Kapur, then the recovery process could be monitored and a gradual return to practice could be responsibly coordinated.

In this current study, even when physicians with scores suggesting cognitive impairment were removed from the competency evaluation group, the performance of that group was lower than a control group of physicians. Thus, it appears that even non-cognitively impaired competency evaluation physicians

may struggle with cognitive shortfalls (Thompson, 2003). For these physicians, the neuropsychological assessment should be done with the motivation to provide them with information about neuropsychological strengths and weaknesses. Compensatory strategies that address cognitive weaknesses could be taught and evaluated for effectiveness. For physicians with specific cognitive deficits, the impact the deficits have on clinical performance would need to be assessed (Kapur, 1997). For example, if a surgeon scores poorly on visual spatial tasks, then his or her ability to continue to perform surgeries is of great concern (Thompson, 2003). A physician health program could help that surgeon consider a career change in medicine that would protect patients and the physician. However, if a psychiatrist scored poorly on the same task, he or she might be able to continue to practice safely, since visual spatial skills may be less critical in psychiatry.

The findings in this current study call into the question the use of non-physician normative groups. While some neuropsychological tests provide for education-related norms, frequently the level of education does not match the high level of education obtained by physicians (Thompson, 2003), as is the case with the MicroCog (Powell et al., 1993). More research on the neurocognitive characteristics of non-impaired physicians is needed. If a physician is compared to groups of people who do not have at least the same level of education, important cognitive deficits could be missed. Thus, a physician who

was tested and told that he or she was functioning in the normal neuropsychological range for the normative group provided by that instrument, could continue to practice and not know of a mild cognitive impairment that could have been addressed, and/or could adversely impact his or her ability to practice safely.

Education plans that are developed for physicians who have completed competency evaluations should incorporate their cognitive characteristics. For example, if a physician has a strong auditory memory, but has difficulty with reading material, then the education plan needs to emphasize auditory learning. Also, compensatory strategies for remembering and processing the large amount of reading material required of physicians would need to be addressed.

Currently, there are few physician normative samples available for neuropsychological tests (Thompson, 2003). Results of this current study indicated that a control group of physicians performed significantly better than a normative sample that had "greater than high school education." Therefore, the development of normative samples that are based on physicians would provide a better comparison group for physicians (Thompson). Based on the experience with this current study, obtaining a normative sample of physicians who are beyond resident status is somewhat difficult. Thus, until these broader samples of cognitively "normal" physicians are available, clinicians should interpret

current test results with the understanding that physicians tend to perform better than a general population or even some education-corrected normative samples.

Future Directions

The issue of physician competency is here to stay. There are many components to consider when addressing this problem and neuropsychological ability needs to be one of the components that is considered. A recent research study at Harvard highlighted the continued concern over physicians' ability to practice safely (Choudhry et al., 2005). In their discussion as to why older physicians are at greater risk for providing lower quality care, there was no mention of the possibility that neuropsychologcial factors could adversely impact older physicians' clinical performance. Thus, it appears that the idea that cognitive deficits are an important piece to understanding physician competency has still not reached full awareness. To facilitate that awareness, more research is needed in a variety of areas.

One area of research that would help make neuropsychological concerns more salient would be in the area of ecological validity of neuropsychological instruments. There are very few studies that provide information about what neuropsychological abilities impact specific physician performance (Thompson, 2003). The more studies that identify the relationship between clinical performance and neuropsychological assessment, the more these assessments will be seen as useful and valid. This information could also be used to provide

research-based counseling for physicians with cognitive deficits. With this type of counseling, physicians could make career decisions with greater confidence (Kapur, 1997).

Future research would benefit from a clear delineation of how cognitive impairment is defined. In reviewing results from the few studies on physicians' neuropsychological characteristics, it was found that very few provided information that operationalized neuropsychological impairment. It is acknowledged that this is no simple task and that part of determining impairment involves, at minimum, a subjective review of the person tested, the pattern of test results, and personal history. However, future research in this area would benefit from development of a common methodology and language for identifying cognitive deficits and impairment. Then results from several studies could be compared and contrasted to provide even more information about the patterns and trends that emerge in this young field.

Much more research in the area of age-related cognitive decline and wisdom, and how these two phenomena interact and impact physician performance, is necessary. It would be helpful to understand what part wisdom, the accumulation of knowledge and experience, plays in compensating for age-related decline, such as processing speed and processing novel information. There would be great benefit from studying successfully practicing older

physicians, their neuropsychological characteristics, and what dynamic efforts they employ to maintain a competent practice.

Summary

In summary, the results from this study suggest that there were significant neuropsychological differences between physicians referred for competency evaluations and physicians whose competency was not in question. The competency evaluation group had a greater proportion of cognitively impaired individuals and performed significantly lower on scores of processing speed, processing accuracy, and cognitive proficiency.

Being a physician is a high and difficult calling. Physicians are asked to make complex, often quick, life and death decisions on a daily basis. A high level of cognitive ability is required for these tasks (Thompson, 2003). This researcher holds great respect for the individuals who have selected this profession. Therefore, it is hoped that these results will be used to assist physicians in pursuing improved clinical performance for all physicians and the identification, evaluation, and potential rehabilitation of impaired physicians.

References

Adams, K. M., & Heaton, R. K. (1987). Computerized neuropsychological assessment: Issues and applications. In Butcher, J. M. (Ed.). Computerized psychological assessment: A practioner's guide (pp. 355-365). NY: Basic Books.

American Medical Association Council on Mental Health. (1973). The sick physician: Impairment by psychiatric disorders, including alcoholism and drug dependence. JAMA, 223, 684-687.

Anastasi, A. (1988). Psychological testing. NY: Macmillan Publishing.

Beaumont, J. G. (1975). The validity of the Category Test administered by on-line computer. Journal of Clinical Psychology,31, 458-462.

Benton, A. L., & Hamsher, K. deS, (1978). Multilingual Aphasia Examination (2nd ed.). Iowa city, IA: University of Iowa Press.

Benton, A. L., Hamsher, K. deS, Varney, N. R., & Spreen, O. (1983). Visual form discrimination. In Contributions to neuropsychological assessment: A clinical manual (pp. 55-62). New York: Oxford University Press.

Branconnier, R. J. (1986). A computerized battery for behavioral assessment in Alzheimer's disease. In L. W. Poon, T. Crook, K. L. Davis, B. J., Eisdorfer, A. W. Gurland, A. W. Kaszniak, & L. W. Thompson (Eds).

Handbook for clinical memory assessment of older adults (pp. 189-196). Washington, D.C.: American Psychological Association.

Browndyke, J. N., Albert, A. L., Malone, W., Schatz, P., Paul, R. H., Cohen, R. A., et al. (2002). Computer-related anxiety: Examining the impact of technology-specific affect on the performance of a computerized neuropsychological assessment measure. Applied Neuropsychology, 9, 210-218.

Butcher, J. N. (Ed.). (1987). Computerized psychological assessment: A practitioner's guide. New York: Basic Books.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). Manual for the restandardized Minnesota Multiphasic Personality Inventory-2: MMPI-2. Minneapolis: University of Minnesota Press.

Butcher, J. N., Perry, J. N., & Atlis, M. M. (2000). Validity and utility of computer-based test interpretation. Psychological Assessment, 12(1), 6-18.

Campbell, K. A., Rohlam, D. S., Storzbach, D., Binder, L. M., Anger, W. K., Kovera, C. A., Davis, K. L., & Grossmann, S. J. (1999). Test-retest reliability of psychological and neurobehavioral tests self-administered by computer. Assessment, 6(1), 21-32.

Cegalis, J.A., & Bowlin, J. (1995). Vigil. For Thought Limited: Nashau, NH.

Center for Personalized Education for Physicians. (n.d.). Retrieved February 28, 2005, from http://cpepdoc.org/.

Chapman, S. (1983). Incompetent physicians: A legal mine field. Physician's Management, 23, 206-207.

Choca, J., & Morris, J. (1992). Administering the Category Test by computer: Equivalence of results. The Clinical Neuropsychologist, 6, 9-15.

Choudhry, M. K., Fletcher, R. H., & Soumeral, S. B. (2005). Systematic review: The relationship between clinical experience and quality of health care. Annals of Internal Medicine, 142, 260-273.

Christensen, H. (2001). What cognitive changes can be expected with normal aging. Australian and New Zealand Journal of Psychiatry, 35, 768-775.

Clay, S. W., & Conatser, R. R. (2003). Characteristics of physicians disciplined by the State Medical Board of Ohio. The Journal of the American Osteopathic Association, 103, 81-88.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155-159.

Collie, A, & Maruff, P. (2003). Computerised neuropsychological testing. British Journal of Sports Medicine, 37(1), 2-3.

Compton, D. M., Bachman, L. D., Brand, D., & Avet, T. L. (2000). Age-associated changes in cognitive function in highly educated adults: Emerging myths and realities. International Journal of Geriatric Psychiatry, 15, 75-85.

Crook, T. H., & Larrabee, G. J. (1988). Interrelationships among everyday memory tests: Stability of factor structure with age. Neuropsychology, 2, 1-12.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). The California Verbal Learning Test. San Antonio, TX: The Psychological Corporation.

Devivo, K., Rothland, J., Price, L., & Fein, G. (1997) Computerized assessment of arithmetic computation skills with MicroCog. Journal of the International Neuropsychological Society, 3, 199-200.

Elkind, J. S., Rubin, E., Rosenthal, S., Skoff, B., & Prather, P. (2001). A simulated reality scenario compared with the computerized Wisconsin Card Sorting Test: An analysis of preliminary results. CyberPsychology & Behavior, 4(4), 489-496.

Elwood, D. L. (1969). Automation of psychological testing. American Psychologist, 24, 287-289.

Elwood, D. L. (1972). Automated versus face-to-face intelligence testing: Comparison of test-retest reliabilities. International Journal of Man-Machine Studies, 4, 363-369.

Elwood, D. L. (1972). Automated WAIS testing correlated with face-to-face intelligence testing: A validity study. International Journal of Man-Machine Studies, 4, 129-137.

Elwood, D. L. (1972). Test-retest reliability and cost analyses of automated and face to face intelligence testing. International Journal of Man-Machine Studies, 4, 1-22.

Elwood, D. L. (1972). Validity of an automated measure of intelligence in borderline retarded subjects. American Journal of Mental Deficiency, 4, 90-94.

Elwood, D. L., & Griffin, H. R. (1972). Individual intelligence testing without the examiner: Reliability of an automated method. Journal of Consulting and Clinical Psychology, 38, 9-14.

Elwood, R. W. (2001). MicroCog: Assessment of Cognitive Functioning. Neuropsychology Review, 11, 89 – 100.

Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. JAMA, 287, 226-235.

Erlanger, D. M., Kaushik, T., Broshek, D., Freeman, J., Feldman, D., & Festa, J. (2002). Development and validation of a web-based screening tool for monitoring cognitive status. Journal of Head Trauma Rehabilitation, 17, 458-476.

Faul, F., & Erdfelder, E. (1992). Gpower: A priori, post-hoc, and compromise power analysis for MS-DOS [Computer Software and Manual]. Retrieved from http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/

Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1997). Computerized neuropsychological tests in the early detection of dementia: Prospective findings. Journal of the International Neuropsychological Society, 3, 139-146.

Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychological applications of CANTAB. International Journal of Geriatric Psychiatry, 11, 329-336.

French, C. C., & Beaumont, J. G. (1990). A clinical study of the automated assessment of intelligence by the Mill Hill Vocabulary Test and the Standard Progressive Matrices Test. Journal of Clinical Psychology, 46, 129-140.

Gendel, J. H. (2000). Disruptive behaviors, personality problems, and boundary violations. In L. S. Goldman, M. Myers, & L. J. Dickstein (Eds.),

Handbook of physician health (pp. 138-160). Chicago: American Medical Association.

Glass, G. V., & Hopkins, K. D. (1996). Statistical methods in education and psychology (3rd ed.). MA: Allyn & Bacon.

Golden, C. J. (1987). Computers in neuropsychology. In Butcher, J. M. (Ed.). Computerized psychological assessment: A practioner's guide (pp. 344-354). NY: Basic Books.

Goldstein, M. Z. (2000). The aging physician. In L. S. Goldman, M. Myers, & L. J. Dickstein (Eds.). Handbook of physician health (pp. 180-192). Chicago: American Medical Association.

Goldstein, G, & Incagnoli, T. M. (Eds.). (1997). Contemporary approaches to neuropsychological assessment. Critical issues in neuropsychology. (pp. 359-392). New York: Plenum Press.

Grantcharov, T. P., Bardram, L., Funch-Jensen, P., & Rosenberg, J. (2001). Laparoscopic performance after one night on call in a surgical department: Prospective study. British Medical Journal,323, 1222-1223.

Green, R. C., Green, J., Harrison, J. M., & Kutner, M. H. (1994). Screening for cognitive impairment in older individuals: Validation of a computer-based test. Archives of Neurology, 51, 779-786.

Gronwall, D. (1977). Paced Auditory Serial-addition task: A measure of recovery from concussion. Perceptual and Motor Skills, 44, 367-373.

Gronwall, D. (1987). Advances in the assessment of attention and information processing after head injury. In H. S. Levin, J. Grafman, & H. M. Eisenberg (Eds.), Neurobehavioral recovery from head injury. New York: Oxford University Press.

Hall, W., Violato, C., Lewkonia, R., Lockyer, J., Fidler, H., Toews, J., Jennett, P., Donoff, M., & Moores, D. (1999). Assessment of physician performance in Alberta: the Physician Achievement Review. Canadian Association Medical Journal, 161, 52 - 57.

Halstead, R. M., & Wolfson, D. (1993). The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation-second edition. Tucson, AZ: Neuropsychological Press.

Halstead, W. C., Walker, A. E., & Bucy, P. A. (1940). Sparing and nonsparing of "macular" vision associated with occipital lobectomy in man. Archives of Ophthalmology, 24, 948-966.

Hanna, E., Premi, J., & Turnbull, J. (2000). Results of remedial continuing medical education in dyscompetent physicians [Electronic version]. Academic Medicine 75, 174-176.

Heaton, R. K. (1993). Wisconsin Card Scoring Test. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., & Adams, K. M. (1987). Potential versus current reality of automation in neuropsychology: Reply to Kleinmuntz. Journal of Consulting and Clinical Psychology, 55, 268-269.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). Wisconsin Card Sorting Manual: Revised and expanded. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). Comprehensive norms for an expanded Halstead-Reitan battery. Odessa, FL: Psychological Assessment Resources.

Heaton, R.K., Thompson, L.L., Nelson, L.M., Filley, CM., & Franklin, G.M. (1990). Brief and Intermediate-length screening of neuropsychological impairment in multiple sclerosis. In Stephen M. Rao (ed.) Neurobehavioral aspects of multiple sclerosis, pp.149-160. New York: Oxford University Press.

Hoffer, P. J., & Green, B. F. (1985). The challenge of competence & creativity in computerized psychological testing. Journal of Consulting and Clinical Psychology, 53, 826-838.

Hughes, P. H., Brandenburg, N., Baldwin, D. C., Jr., Storr, C. L., Williams, K. M., Anthony, J. C., & Sheehan, D. V. (1992). Prevalence of substance use among US physicians. JAMA, 267, 2333 - 2339.

Jacobs, D. M., Sano, M., Albert, S., Schofield, P., Dooneief, G., & Stern, Y. (1997). Cross-cultural neuropsychological assessment: A comparison of randomly selected, demographically matched cohorts of English- and Spanish-speaking older adults. Journal of clinical and Experimental Neuropsychology, 19, 331-339.

Johnson, D. F., & Mihal, W. L. (1973). Performance of blacks and whites in computerized versus manual testing environments. American Psychologist, 28, 694-699.

Johnson, D. F., & White, C. B. (1980). Effects of training on computerized test performance in the elderly. Journal of Applied Psychology, 65, 357-358.

Kane, R. L. (1998). MicroCog: A review. NAN Bulletin, 11, 13-16.

Kane, R. L., & Kay, G. G. (1992). Computerized assessment in neuropsychology: A review of tests and test batteries. Neuropsychology Review, 3(1), 1 – 117.

Kane, R. L., & Kay, G. G. (1997). Computer applications in neuropsychological assessment. In G. Goldstein, & T. M. Incagnoli, (Eds.) Contemporary

approaches to neuropsychological assessment. Critical issues in neuropsychology (pp. 359-392). New York: Plenum Press.

Kane, R. L., & Reeves, D. L. (1997). Computerized test batteries. In A. M. Horton, Jr., D. Wedding, & J. Webster, (Eds.). The neuropsychology handbook, volume 1: Foundations and assessment (2nd ed., pp. 423-467). New York: Springer Publishing.

Kapur, N. (Ed.). (1997). Injured brains of medical minds. Oxford: Oxford University Press.

Kawas, D., Gray, S., Brookmeyer, R., Fozard, J., & Zonderman, A. (2000). Age-specific incidence rates of Alzheimer's disease: The Baltimore Longitudinal Study of Aging. Neurology, 54, 2072-2077.

Knights, R. M., & Watson, P. (1968). The use of computerized test profiles in neuropsychological assessment. Journal of Learning Disabilities, 1, 696-709.

Knights, R., Richardson, D., & McNarry, L. (1973). Automated vs. clinical administration of the Peabody Picture Vocabulary Test and the Raven's Coloured Progressive Matrices. American Journal of Mental Deficiency, 78, 223-225.

Lee, J. (1986). The effects of past computer experience on computerized aptitude test performance. Educational and Psychological Measurement, 46, 727-733.

Lewis, S. B. (1983). The physically handicapped physician. In J. P. Callan (Ed.), The physician: A professional under stress. (pp. 318-326). East Norwalk, Connecticut: Appleton-Centruy-Crofts.

Long, C. J., & Wagner, M., (1986). Computer applications in neuropsychology. In D. Wedding, A. M. Horton, & J. Webster (Eds.), The neuropsychology handbook: Behavioral and clinical perspectives. (pp. 548-569). NY: Springer Publishing.

Lopez, S. J. (1999). Validity of the MicroCog: Assessment of Cognitive Functioning in a clinical sample of substance abusers. Dissertation Abstracts International 59(12-B), 6515. (UMI Number 9914103)

Lunsford, T. E. (1981). Problem of the aging physician: A report of the commission on physician impairment. Journal of the Indiana State Medical Association, 74, 778-779.

Madden, D. J. (1988). Cognitive impairment in physicians. Maryland Medical Journal, 37, 201-205.

Maddox, G. L. (1987). Psychosocial perspectives of aging. In M. W. Riley, & J. D. Matarazzo, The aging dimension. Perspectives in behavioral medicine (pp. 15-28). Hillsdale, NJ: Lawrence Erlbaum Associates

Martini, C. J. M. (1987). Physical disabilities and the study and practice of medicine. JAMA, 257, 2956-2957.

Matarazzo, J.D., & Goldstein, S. G. (1972). The intellectual caliber of medical students. Journal of Medical Education, 47, 102.111.

Mattis, S. (1973, 1988). Dementia Rating Scale. Odessa, FL: Psychological Assessment Resources.

Maulucci, R. A., & Eckhouse, H. R. Jr. (1988). The use of computers in the assessment and treatment of cognitive disabilities in the elderly: A survey. Psychopharmacology Bulletin, 24, 557-564.

McAuley, R. G., & Henderson, H. W. (1984). Results of the peer assessment program of the College of Physicians and Surgeons of Ontario. Canadian Medical Association Journal, 131, 557-561.

McAuley, R. G., Paul, W. M., Morrison, G. H., Beckett, R. F., & Goldsmith, C. H. (1990). Five year results of the peer assessment program of the College of Physicians and Surgeons of Ontario. Canadian Medical Association Journal, 43, 1193-1199.

McAuliffe, W. E., Rohman, M., Breer, P., Wyshak, G., Santangelo, S., & Magnuson, E. (1991). Alcohol use and abuse in random smaples of physicians and medical students. American Journal of Public Health, 81, 177-182.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). EDITS Manuel for the Profile of Mood States. San Diego, CA: Educational and Industrial Service.

Meehl, P. E. (1954, 1996). Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Northvale, NJ: Jason Aronson.

Mehiltz, M. Kleinoedr, T, Weniger, G., & Rienhoff, O. (1998). Design of a virtual reality laboratory for interdisciplinary medical application. Medinfo, 9, Pt 2: 1051-1055.

Moore, R. D., Mead, L., & Pearson, T. A. (1990). Youthful precursors of alcohol abuse in physicians. American Journal of Medicine, 88, 332-336.

Morrison, J., & Wickersham, P. (1998). Physicians disciplined by a state medical board. JAMA, 279, 1889-1893.

Murray, D., & Dodds, C. (2003). The effect of sleep disruption on performance of anaesthetists - A pilot study. Anaesthsia, 58, 520-526.

Nelson, E. A., & Dannefer, D. (1992). Aged heterogeneity: Fact or fiction? The fate of diversity in gerontological research. Gerontologist, 32, 17-23.

Nelson, J. H. (1999). The efficacy of MicroCog as an outcome measure of gains in cognitive rehabilitation. Dissertation Abstracts International 59(8-B), 4477. (UMI No. 9725266)

Nixon, S. J. (1996). Secondary Dementias: Reversible Dementias and Pseudodementia. In R. L. Adams, O. A. Parsons, J. L. Culbertson, & S. J. Nixon (Eds.), Neuropsychology for clinical practice: Etiology, assessment, and treatment of common neurological disorders (pp. 107-130). Washington DC: American Psychological Association.

Norcini, J. J., Lipner, R. S., Benson, J. A., Jr., & Webster, G. D. (1985). An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. Annals of Internal Medicine, 102, 385-389.

Norman, G. R., Davis, D. A., Lamb, S., Hanna, E., Caulford, P., & Kaigas, T. (1993). Competency assessment of primary care physicians as part of a peer review program. JAMA, 270, 1046-1052

O'Keefe, K. E. (1997). Validation of a computerized neuropsychological test. Dissertation Abstracts International 58(3-B), 1589. (UMI No.9725266)

Overton, G. W., & Scott, K. G. (1972). Automated and manual intelligence testing: Data on parallel forms of the Peabody Picture Vocabulary Test. American Journal of Mental Deficiency, 76, 639-643.

Pallant, J. (2001). SPSS survival manual: A step by step guide to data analysis using SPSS for Windows (Verions 10 and 11). New York: Open University Press.

Pasko, T., & Seidman, B. (2002). Physician characteristics and distribution in the US: 2002-2003 Edition. Chicago: American Medical Association.

Pellegrino, J. W., Hunt, E. B., Abate, R., & Farr, S. (1987). A computer-based test battery for the assessment of static and dynamic spatial reasoning abilities. Behavior Research Method, Instruments, & Computers, 19, 231-236.

Perry, W., & Crean, R. D. (2005). A retrospective review of the neuropsychological performance of physicians referred for medical infractions. Archives of Clinical Neuropsychology, 20, 161-170.

Powell, D. H. (1997). Comment on computerized assessment of arithmetic computation skills with MicroCog. Journal of the International Neuropsychological Society, 3, 200.

Powell, D.H., Kaplan, E.F., Whitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. H. (1993). MicroCog assessment of Cognitive Functioning. San Antonio, TX: The Psychological Corporation.

Powell, D. H., & Whitla, D. K. (1994). Profiles in cognitive aging. Cambridge, MA: Harvard University Press.

Pugnetti, L., Mendozzi, L., Notta, A., Cattaneo, A., Barbieri, E., & Brancotti, A. (1995). Evaluation and retraining of adults' cognitive impairment: Which role for virtual reality technology? Computer Biological Medicine, 25, 213-227.

Raven, J. C., Court, J. H., & Raven, J. (1986). Raven's Progressive Matrices. London: H. K. Lewis.

Regier, C. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L., & Goodwin, F. K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse: Results from the Epidemiologic Catchment Area study. JAMA, 264, 2511-1518.

Reich, D. L., Uysal, S., Bodian, C. A., Gabriele, S., Hibbard, M., Gordon, W. et al. (1999). The relationship of cognitive, personality, and academic measures to anesthesiology resident clinical performance. Anesthesia and Analgesia, 88, 1092-1100.

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. Perceptual and Motor Skills, 8, 271-276.

Reitan, R. M., & Wolfson, P.(1985). The Halstead-Reitan Neuropsychological Test Battery; Theory and clinical interpretation, Tucson: Neuropsychology Press.

Reitan, R. M., & Wolfson, P.(1993). The Halstead-Reitan Neuropsychological Test Battery; Theory and clinical interpretation (2nd ed.), Tucson: Neuropsychology Press.

Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. Archives de Psychologie, 28, 286-340.

Rey, A. (1959). Sollicitation de la memoire de fixation par des mots et des objets presentes simultanement. Archives de Psychologie, 37, 126-139.

Rey, A. (1964). L'esamen clinique en psycholgie. Paris: Presses Universitaires de France.

Ritchie, K. A., & Hallerman, E. F. (1989). Cross-validation of a dementia screening test in a heterogeneous population. International Journal of Epidemiology, 18, 717-719.

Robbins, J., & Gottlieb, F. (1990). Sleep deprivation and cognitive testing in internal medicine house staff. The Western Journal of Medicine, 152, 82-86.

Roid, G. H. (1986). Computer technology in testing. In B. S. Plake, & J. C. Witt, Buros-Nebraska Symposium on measurement and testing. Volume 2: The future of testing. (pp. 29-69). Hillsdale, NJ: Lawrence Erlbaum.

Rorschach, H. (1921). Psychodiagnostik. Bern: Bircher (Transl. Hans Huber Verlag, 1942).

Russell, E. W., Neuringer, C., & Goldstein, G. (1970). Assessment of brain damage: A neuropsychological key approach. New York: Wiley-Interscience.

Sadavoy, J. (1994). The aging physician. Journal of Geriatric Psychiatry, 27, 265-279.

Sargent, D. A. (1985). The impaired physician movement: An interim report. Hospital and Community Psychiatry, 36, 294-297.

Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. Psychology and Aging, 8, 44-55.

Schatz, P., & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. Journal of Head Trauma Rehabilitation: Special Issue: Neuropsychological Technologies, 17, 395-410.

Schatz, P., & Zillmer, E. A. (2003). Computer-based assessment of sports-related concussion. Applied Neuropsychology, 10(1), 42-47.

Schueneman, A. R., Pickleman, J., Hesslein, R., & Freeark, R. J. (1984). Neuropsychologic predictors of operative skill among general surgery residents. Surgery, 97, 288 – 295.

Shipley, W. C. (1946). Shipley Institute of Living Scale. Los Angeles, CA: Western Psychological services.

Smith, A. (1973) Symbol Digit Modalities Test. Los Angeles, CA: Western Psychological Services.

Smith, R. E. (1963). Examination by computer. Computers in Behavioral Science, 8, 76-79.

Smith-Coggins, R., Rosekind, M., Hurd, S., & Buccino, K. (1994). Relationship of day versus night sleep to physician performance and mood. Annals of Emergency Medicine, 24, 928-934.

Space, L. G. (1981). The computer as psychometrician. Behavaior Research Methods & Instrumentation, 13. 595-606.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662.

Stulp de Groot, C., Barely, G., Dickerman, J., & Korinek, E. (2004). Reasons physicians are referred for personalized assessment and education: A CPEP report. Manuscript submitted for publication.

Tabachnick, B. G., & Fidell, L. S. (1996). Using multivariate statistics. Boston, MA: Allyn & Bacon.

Taffinder, N. J., McManus, I. C., Gul, Y., Russell, R. C., & Darzi, A. (1998). Effect of sleep deprivation on surgeons' dexterity on laparoscopy simulator. Lancet, 352, 1191.

Terry, K. (2002). Impaired physicians: Speak no evil? Medical Economics, 79, 110-112, 114, &117.

Thompson, L. L. (2003). Neuropsychological assessment of physicians whose competency to practice medicine is being questioned. In G. P. Prigatano, & M. H. Pliskin (Eds.), Clinical Neuropsychology and Cost Outcome Research: A Beginning. New York: Psychology Press.

Trunkey, D. D., & Botney, R. (2001). Assessing competency: A tale of two professions. Journal of the American College of Surgeons, 192, 385-395.

Turnbull, J., Carbotte, R., Hanna, E., Norman, G., Cunnington, J., Ferguson, B., & Kaigas, T. (2000). Cognitive difficulty in physicians [Electronic version]. Academic Medicine, 75, 177-181.

Van Komen, G. J. (2000). Troubled or troubling physicians: Administrative responses. In L. S. Goldman, M. Myers, & L. J. Dickstein (Eds.), The handbook of physician health: The essential guide to understanding the

health care needs of physicians (pp. 205-226). Chicago: American Medical Association.

van Zomeren, A. H., & Brouwer, W. H. (1992). Assessment of attention. In J. R. Crawford, D. M. Parker (Eds.), A handbook of neuropsychological assessment (pp. 241-266). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Watts, K., Baddeley, A. D., & Williams, M. (1982). Automated tailored testing using Raven's matrices and the Mill Hill Vocabulary tests: A comparison with manual administration. International Journal of Man-Machine Studies, 17, 331-344.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale-Revised. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1987). Wechsler Memory Scale-Revised. San Antonio, TX: The Psychological Corporation.

Weinger, M. B., & Ancoli-Israel, S. (2002). Sleep deprivation and clinical performance. JAMA, 287, 955-957.

Weiss, D. J., & Vale, C. D. (1987). Computerized adaptive testing for measuring abilities and other psychological variables. In Butcher, J. M. (Ed.). (1987). Computerized psychological assessment: A practioner's guide (pp. 325-343). NY: Basic Books.

Wesnes, K., Walker, M., & Walker, L. (1997). Cognitive performance and mood after a weekend on call in a surgical unit. British Journal of Surgery, 84, 493-495.

Wetzel, L. & Boll, T. J. (1987). Short Category Test, Booklet Format. Los Angeles, CA: Western Psychological Services.

Wetzler, S. (1990). Computerized psychological assessment. In Baskin, D. (Ed.) Computer applications in psychiatry and psychology (pp. 43-56.). New York: Brunner/Mazel.

White, R. K., & Hayes, M. G. (1987). The physician rehabilitation Program of Maryland. Maryland Medical Journal, 36, 221-224.

White, R. F., James, K. E., Vasterling, J. J., Letz, R., Marans, K., Delaney, R., et al. (2003). Neuropsychological screening for cognitive impairment using computer-assisted tasks. Assessment, 10, 86-101.

Wilkinson, G. S. (1993). Wide Range Achievement Test-Third edition. Wilmington, DE: Jastak Associates.

Williams, B. W., Williams, M. V., & Norcross, W. A. (2002, October). Differences in cognitive performance between disciplined and non-disciplined physicians and foreign versus domestic medical schools. Paper presented at the meeting of the North American Primary Care Research Group Annual Meeting, New Orleans, LA.

Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., Evans, D. A., & Bennett, D. A. (2002). Individual differences in rates of change in cognitive abilities of older persons. Psychology and Aging, 17, 179-193.

Wingenfeld, S. A., Holdwick, D. J., Jr., Davis, J. L., & Hunter, B. B. (1999). Normative data on computerized paced auditory serial addition task performance. Clinical Neuropsychologist, 13, 268-273.

Zachary, R. A. (1992) Shipley Institute of Living Scale-Revised. Los Angeles, CA: Western Psychological services.

Zelinski, E. M., & Kennison, R. F. (2001). The Long Beach Longitudinal Study: Evaluation of longitudinal effects of aging on memory and cognition. Home Health Care Services Quarterly, 19, 45-55.

APPENDIX A

Institutional Review Board Approvals and Continuation Approvals

University of Denver Institutional Review Board - Approval

UNIVERSITY OF
DENVER

Office of Sponsored Programs
2199 S. University Blvd.
Denver, CO 80208
303.871.2121
Fax 303.871.2623

May 26, 2004

Ms. Lori Korinek
3469 South Lake Gulch Road
Castle Rock, CO 80104

Subject: Human Subject Review
TITLE: "Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations."
IRB: 04044
PURPOSE: Ph.D. Dissertation
SPONSOR: None
NEW APPLICATION

Dear Ms. Korinek:

The Institutional Review Board for the Protection of Human Subjects has reviewed the above named project. The project has been approved for the procedures and subjects described in the protocol at the **April 13, 2004** meeting. This approval is effective for twelve months. We will be sending you a continuation application for this project in **December, 2004.** This form must be completed and returned to the Office of Sponsored Programs if the project is to be continued. If you do not receive this application, please contact Dawn Nowak, dnowak@du.edu.

Attached is a copy of your consent form, as approved by the Institutional Review Board, a brief summary of your responsibilities regarding the use of human subjects, and a copy of the University of Denver Assurance of Compliance with Health and Human Services Regulations for Protection of Human Subjects.

The Institutional Review Board appreciates your cooperation in protecting subjects and ensuring that each subject gives a meaningful consent to participate in research projects. If you have any questions regarding your obligations under the Assurance, please do not hesitate to call Dawn Nowak.

Sincerely yours,

Dr Norman Watt/dn

Dr. Norman Watt, Ph.D.
Acting Chair, Institutional Review Board
for the Protection of Human Subjects

Attachments

cc: C. McCrae

Colorado Multiple Institutional Review Board - Approval

Colorado Multiple Institutional Review Board
13001 E. 17th Place
Building 500, Room N3214
Aurora, Colorado 80010-7238

Mailing Address:
Mail Stop F-490
P.O. Box 6508
Aurora, Colorado 80045-6508

(303) 724-1055 [Phone]
(303) 724-0990 [Fax]
www.uchsc.edu/comirb [Web]
comirb@uchsc.edu [E-Mail]
FWA#: FWA00005070

University of Colorado Hospital
Denver Health Medical Center
Veteran's Administration Medical Center
The Children's Hospital
University of Colorado Health Sciences Center
Colorado Prevention Center

10/07/2004

Certificate of Approval

Investigator: Lauri Korinek

Sponsor(s):

Subject: **COMIRB Protocol 04-0610**
Initial Review (APP001)
2nd

Title:

THE NEUROPSYCHOLOGICAL DIFFERENCES BETWEEN A CONTROL GROUP OF PHYSICIANS AND PHYSICIANS REFERRED FOR COMPETENCY EVALUATIONS.

Approval Date: **6 October 2004**

Expiration Date: **6 October 2005**

Approval Includes: **Protocol - investigator - 1 Consent and/or Assent Form(s) - 4 Advertisement(s) - 1 Questionnaire(s)**

All COMIRB Approved investigators must comply with the following:

- For the duration of your protocol, any change in the experimental design/consent and/or assent form must be approved by the comirb before implementation of the changes.
- Use only a copy of the **COMIRB signed and dated Consent and/or Assent Form.** The investigator bears the responsibility for obtaining from all subjects "Informed Consent" as approved by the COMIRB. The COMIRB REQUIRES that the subject be given a copy of the consent and/or assent form. Consent and/or assent forms must include the name and telephone number of the investigator.
- **Provide non-English speaking subjects with a certified translation of the approved Consent and/or Assent Form** in the subject's first language. A copy of the translator's certification should be attached to the consent and/or assent form.
- The investigator also bears the responsibility for **informing the COMIRB immediately of any Serious Adverse Events** (deaths, serious complications or other untoward effects of this research at this or other sites), and of the relationship of the SAE to the investigational trial. The COMIRB uses the standard definition of Serious or Unanticipated Events that inclue: death, hospitalization, prolongation of hospitalization and other unanticipated side effects
- Obtain COMIRB **approval for all advertisements** before use.
- **Federal regulations require a Continuing Review to renew approval of this project** within a 12-month period from the last approval date unless otherwise indicated in the review cycle listed below. If you have a restricted/high risk protocol, specific details will be outlined in this letter. Non-compliance with Continuing Review will result in the termination of this study. This project has been assigned the following review cycle:

COMIRB Continuing Review Cycle: 12 months

We will send you a Continuing Review Form to be completed prior to the due date. Any questions regarding the COMIRB action on this study should be referred to the COMIRB staff at 303-724-1055 or UCHSC Box F-490.

Norman Stoller / Jamie Hoey

Ken Easterday, RPh, Co-Chair
Christopher Kuni, MD, Co-Chair
Norman Stoller, DMD, Co-Chair
Hans Neville, MD, Co-Chair
Dave Lawellin, PhD, Co-Chair
Douglas Ford, MD, Co-Chair
Steve Bartlett, RPh, Co-Chair

Revised 09/04

04-0610 Panel: X/B Expedited

University of Denver Continuation Approval

UNIVERSITY OF
DENVER

Office of Sponsored Programs
2199 S. University Blvd.
Denver, CO 80208
303.871.2121
Fax 303.871.2623

March 11, 2005

Ms. Lori Korinek
3469 South Lake Gulch Road
Castle Rock, CO 80104

Subject: Human Subject Review
TITLE: "Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations."
IRB: 04044
PURPOSE: Ph.D. Dissertation
SPONSOR: None
CONTINUATION APPLICATION

Dear Ms. Korinek:

The Institutional Review Board for the Protection of Human Subjects has reviewed the above named project. The project has been approved for the procedures and subjects described in the protocol at the **March 8, 2005** meeting. This approval is effective for twelve months. We will be sending you a continuation application for this project in **November, 2005.** This form must be completed and returned to the Office of Sponsored Programs if the project is to be continued. If you do not receive this application, please contact Dawn Nowak, dnowak@du.edu.

Attached are a copy of your consent form, as approved by the Institutional Review Board, a brief summary of your responsibilities regarding the use of human subjects, and a copy of the University Of Denver Assurance Of Compliance with Health and Human Services Regulations for Protection of Human Subjects.

The Institutional Review Board appreciates your cooperation in protecting subjects and ensuring that each subject gives a meaningful consent to participate in research projects. If you have any questions regarding your obligations under the Assurance, please do not hesitate to call Dawn Nowak.

Sincerely yours,

Dr. Stephen Shirk /DN

Dr. Stephen Shirk, Ph.D.
Acting Chair, Institutional Review Board
For the Protection of Human Subjects

Attachments
cc: C. McRae

APPENDIX B

HIPPA Forms

IRB Application – L. L. Korinek, M.A.

WAIVER OF HIPAA AUTHORIZATION	DU IRB Number: 040454
Protected Health Information is defined as "Individually identifiable health information that a health care provider, health plan, health care clearinghouse or employer creates or receives and includes information about the past, present or future physical or mental health of a person, the provision of health care to a person or the payment for the provision of care to that person."	
1. Describe the protected health information (PHI) that will be collected: **None. This is from an existing data set, without identifying data. Information including age, specialty, nine index scores from the MicroCog, reason for referral, gender, and handedness will be collected from individual files and placed in a database. No one is older than 80 years old.**	
IN ORDER FOR THIS WAIVER TO BE APPROVED, THERE MUST BE NO MORE THAN MINIMAL RISK TO PRIVACY OF THE SUBJECT, BASED ON THE ANSWERS TO THE FOLLOWING QUESTIONS:	
2a. How will subject identifiers be protected? **All identifying information has been removed.**	
2b. What is the plan to destroy the identifiers ASAP? [Please state if there is a health or research justification for retaining the identifiers of if retention is required by law.] **No identifiers are currently attached to data set.**	
2c. Will the data be made available to anyone other than the study personnel? If so, to whom? And if so, why? **No**	

	YES	NO
3. Can this project be done without PHI?	**x**	
4. Why is it not possible to get the authorization of the subjects whose PHI you want to use? **It is not possible because the identifying information has been removed from the files prior to the time they became available for my use.**		

CONFIRMATION:
I confirm that the Protected Health Information (PHI) will not be re-used or disclosed except as required by law, for authorized oversight of the research or for other research that has been reviewed and approved by the IRB with specific approval regarding access to this PHI.

MA 3/22/04

PI Signature Date

FOR DU IRB USE ONLY:

This protocol has been reviewed at a meeting of the IRB and:

____ meets waiver criteria . Alteration, or waiver, in whole or in part, of authorization has been satisfied by the presence of the above criteria and is granted.
____ does not meet the criteria for approval of Waiver of Authorization

Signature of DU IRB Chair Date

WAIVER OF HIPAA AUTHORIZATION	**COMIRB Number:** **PI Name: Lauri Korinek**
Protected Health Information is defined as "Individually identifiable health information that a health care provider, health plan, health care clearinghouse or employer creates or receives and includes information about the past, present or future physical or mental health of a person, the provision of health care to a person or the payment for the provision of care to that person."	
1. Check the protected health information (PHI) that will be used: ☐ Names ☐ Geographic subdivisions smaller than a state ☐ All dates (except year) that are directly related to an individual (e.g. DOB, discharge date) ☐ Telephone numbers ☐ Fax numbers ☐ Electronic mail addresses ☐ Social security numbers ☐ Medical record numbers ☐ Health plan beneficiary numbers ☐ Account numbers ☐ Certificate/license numbers ☐ Vehicle identifiers and serial numbers ☐ Device identifiers and serial numbers ☐ URLs (http://....) ☐ IP address numbers ☐ Biometric identifiers (including finger and voice prints) ☐ Full face photographic images and any comparable images ☐ Any other unique identifying number, characteristic or code **2. Describe where the specific PHI will be accessed from:** No identifying information is part of this data set.	
3. Describe any contact the study personnel will have with the subject: (i.e. office visits, phone calls, mailings, e-mails, etc) None	
IN ORDER FOR THIS WAIVER TO BE APPROVED, THERE MUST BE NO MORE THAN MINIMAL RISK TO PRIVACY OF THE SUBJECT, BASED ON THE ANSWERS TO THE FOLLOWING QUESTIONS:	
4a. How will subject identifiers be protected? There are no identifiers in the data set.	
4b. What is the plan to destroy the identifiers ASAP? [Please state if there is a health or research justification for retaining the identifiers of if retention is required by law.] There are no identifiers for this data set.	

4c. Will the data be made available to anyone other than the study personnel? If so, to whom? And if so, why?

No

	YES	NO
5. Can this project be done without PHI?	X	

6. Why is it not possible to get the authorization of the subjects whose PHI you want to use?

No identifying information is available to obtain authorization.

Attach a copy of the currently approved protocol summary to this form with submission.

CONFIRMATION:

I confirm that the Protected Health Information (PHI) will not be re-used or disclosed except as required by law, for authorized oversight of the research or for other research that has been reviewed and approved by the IRB with specific approval regarding access to this PHI.

PI Signature — Date: 9/8/04

FOR COMIRB USE ONLY:

This protocol has been reviewed under expedited procedures and the use and disclosure of PHI and:

___ meets the above criteria . Alteration, or waiver, in whole or in part, of authorization has been satisfied by the presence of the above criteria and is granted.
____ does not meet the criteria for approval of Waiver of Authorization

Signature of COMIRB Chair — Date: 9/9/04

☐ **Steve Bartlett, R.Ph.** ☐ **Ken Easterday, R.Ph.** ☐ **David Lawellin, Ph.D.**
☐ **Hans Neville, M.D.** ☐ **Cornelis Rietmeijer, M.D.** ☐ **Doug Ford, M.D.**
☑ **Norman H. Stoller, DMD**

HIPPA B Form - COMIRB Participants Only

Authorization To Use or Release Health Information About Me For Research Purposes *Authorization B: Enrollment into Research*	Study Title: Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations. COMIRB Number: 04-0610

I _____ *(Subject's Full Name)*
authorize

Lauri L. Korinek, M.A. *(PI or Physician Name)* **and staff members of**

One research assistant from DU *(Facility Name)* **working for him/her to use the following health information about me <u>for research</u>:** *(Please check the appropriate boxes. NOTE: If a category is checked "yes" and a line follows the category, you MUST describe the type of the procedures done.)*

No Yes
☐ ☒ Name and/or phone number
☐ ☒ Demographic information (age, sex, ethnicity, address, etc.)
☐ ☐ Diagnosis(es)
☐ ☐ History and/or Physical
☐ ☐ Laboratory or Tissue Studies: _____
☐ ☐ Radiology Studies: _____
☐ ☐ Testing for or Infection with Human Immunodeficiency Virus (HIV) (or results)
☐ ☐ Procedure results: _____
☐ ☒ Psychological tests: MicroCog: Assessment of Cognitive Abilities
☐ ☐ Survey/Questionnaire: _____
☐ ☐ Research Visit records
☐ ☐ Portions of previous Medical Records that are relevant to this study
☐ ☐ Billing or financial information
☐ ☐ Drug Abuse
☐ ☐ Alcoholism or Alcohol abuse
☐ ☐ Sickle Cell Anemia
☐ ☐ Other (Specify): _____

For the Specific Purpose of
☒ Collecting data for this research project
☐ Other* N/A

**Cannot say "for any and all research", "for any purpose", etc.*

If my health information that identifies me is also going to be given out to others outside the facility, the recipients are described on the next page(s).
☒ **No personally identifiable health information about me will be disclosed to others**

The PI (or staff acting on behalf of the PI) will also make the following health information about me available to: (check all that apply and **describe the type of the procedures** done where applicable)

Recipient (name of person or group)

No Yes

☐ ☐ All Research Data Collected in this Study (if you check this box Yes, no other boxes need to be checked in this section)

☐ ☐ Name and phone number
☐ ☐ Demographic information (age, sex, ethnicity, address, etc.)
☐ ☐ Diagnosis(es)
☐ ☐ History and Physical
☐ ☐ Laboratory or Tissue Studies: _____
☐ ☐ Radiology Studies: _____
☐ ☐ Testing for or Infection with Human Immunodeficiency Virus (HIV) (or results)
☐ ☐ Procedure results: _____
☐ ☐ Psychological tests: _____
☐ ☐ Questionnaire/Survey: _____
☐ ☐ Research Visit records
☐ ☐ Portions of previous Medical Records that are relevant to this study
☐ ☐ Billing/Charges
☐ ☐ Drug Abuse
☐ ☐ Alcoholism or Alcohol
☐ ☐ Sickle Cell Anemia
☐ ☐ Other (Specify): _____

For the Specific Purpose of

☐ Evaluation of this research project
☐ Evaluation of laboratory/tissue samples
☐ Data management
☐ Data analysis
☐ Other*: _____

**Cannot say "for any and all research", "for any purpose", etc.*

For additional Recipients, copy this page as needed.

I give my authorization knowing that:

- I do not have to sign this authorization. But if I do not sign it the researcher has the right to not let me be in the research study.
- I can cancel this authorization any time.
  - I have to cancel it in writing.
  - If I cancel it, the researchers and the people the information was given to will still be able to use it because I had given them my permission, but they won't get any more information about me.
  - If I cancel my authorization, I may no longer be able to be in the study.
  - I can read the Notice of Privacy Practices at the facility where the research is being conducted to find out how to cancel my authorization.
- The records given out to other people may be given out by them and might no longer be protected.
- I will be given a copy of this form after I have signed and dated it.

This authorization will expire on: _____ (Date) OR

☒ The end of the research study

☐ Will not expire

☐ _____

(Describe dates or circumstances under which the authorization will expire.)

ADDITIONAL INFORMATION: Identification numbers will be used on all paperwork. Your name and identification number will be kept on a list for 30 days after your test date to allow for individual feedback. Upon completion of individual feedback or 30 days after your test date (which ever comes first) your name will be removed from that list and we will be unable to connect your name with your test results. All identifying information will be destroyed at the end of the study.

_____

Subject's Signature Date

_____

Signature of Legal Representative (If applicable) Date

_____

Name of Legal Representative (please print)

_____

Description of Legal Authority to Act on Behalf of Patient

APPENDIX C

Demographic Form

Demographic Form - Participant Number_____

1. Gender: M F
2. Race: (Please mark all that apply.)
   _____White
   _____Black, African American
   _____American Indian or Alaska Native
   _____Asian
   _____Native Hawaiian and Other Pacific Islander
   _____Other, please specify_____

3. Are you Spanish/Hispanic/Latino? (Circle one) Yes No
   If yes, please print group._____

4. Age:_____

5. Please indicate your current job status: (Check One)
   _____Full time
   _____Part time
   _____Retired/Volunteering in the Medical Field

6. How many years have you been practicing medicine, post-residency?_____
   years

7. Please write in your specialty area._____

8. Please indicate if you have had any known neuropsychological related illnesses or accidents that involved a head injury. (Indicate all that apply.)
   _____Seizures
   _____Traumatic Brain Injury
   _____CVA (Cerebrovascular Accident)
   _____Tumor
   _____Movement Disorder
   _____Other Neurological Insults
   _____Other Diseases that Impact Neurological Functioning

9. Please indicate your experience with computers.
   _____Very experienced - Use them every day
   _____Moderate experience - Use them more than once a week
   _____Some experience - Use them less than once a week
   _____No experience - Don't know how to turn one on

APPENDIX D

Business Card

APPENDIX E

Initial Email Advertisement

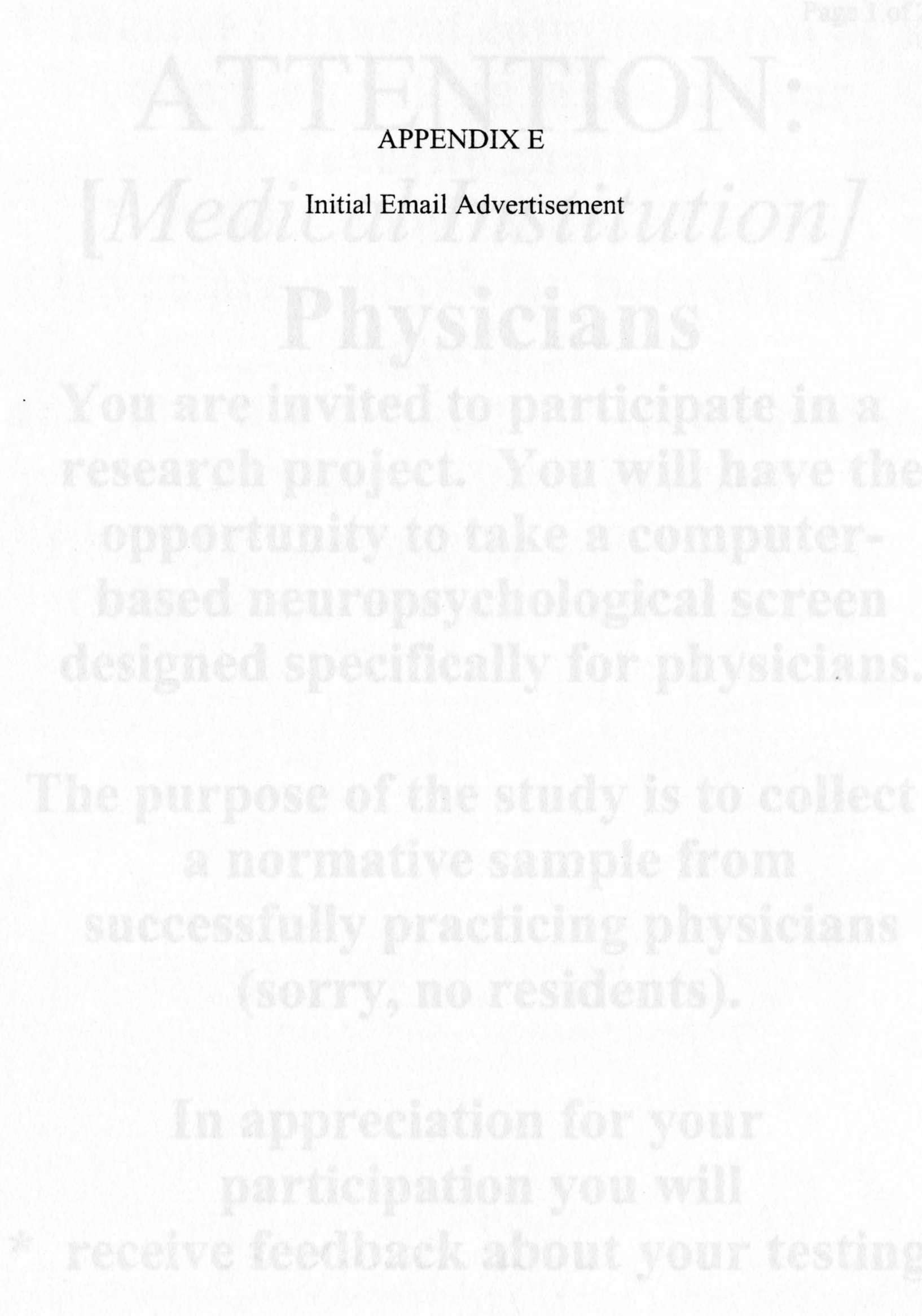
You Are Invited to Participate
in a Research Study About
The Neuropsychological Characteristics of
Physicians

Interested?

Call: Lauri Korinek 303-324-0451

APPENDIX E

Initial Email Advertisement



ATTENTION:

[*Medical Institution*]

Physicians

You are invited to participate in a research project. You will have the opportunity to take a computer-based neuropsychological screen designed specifically for physicians.

The purpose of the study is to collect a normative sample from successfully practicing physicians (sorry, no residents).

In appreciation for your participation you will

*** receive feedback about your testing**

*** receive financial compensation or a donation will be made to your favorite charity**

*** receive food - a boxed meal**

If you are a physician between the ages of 30 and 80 years old and are interested in the study, please contact Lauri L. Korinek at 303-324-0451. Or Read On For More Information

Dear (Medical Institution) Physician,

You are invited to participate in a study to collect normative data on a screen of neuropsychological functioning that was designed specifically for physicians. The MicroCog (Powell et al., 1993), a computer-based assessment, was originally developed to assess cognitive functioning of physicians. However, due to changes in the original test, no current normative sample of physicians exists. In addition to its previous uses, we hope that this test will be used to identify and encourage physicians to work beyond the normal retirement age. This will be important in the next few decades as the large number of "baby-boomer" physicians retire.

A second purpose of this study is to gather data from a group of physicians who are successfully practicing medicine and are 30 to 80 years old. We intend to use this information to better understand the cognitive picture of physicians who are practicing successfully in comparison with those who have been involved in competency evaluations.

Participation in this study will involve a one-hour testing session with a research assistant or this investigator. You will receive general feedback about your test results, with the option of receiving more specific feedback upon request. In appreciation for your participation in the study, you will either receive $50 or a donation of $50 will be made to a charity of your choice. Food will also be provided.

All information gathered for this study is confidential. An identification number will be used on all paperwork. Only the researcher will have the list that matches this number with your name, and this list will be kept in an undisclosed secure setting. In addition, when the researcher reports information, it will be reported for the entire group of participants. Individual results will only be reported to the respective participant. Once individual feedback is provided your name will be removed from the list matching your identification number and your name. Upon completion of data collection and feedback, all identifying data will be destroyed.

I hope you will consider participating in this valuable study. Please respond to this e-mail (via telephone or e-mail) regardless of whether you are interested in participating in this study or not, to eliminate future requests for participation. If you are interested in participating in this study, please include a phone number with your response and you will receive a call. Further information

about the study will be provided, and an appointment will be scheduled if you choose to participate. If you have further questions, please contact Lauri L. Korinek at 303-324-0451 or at LLKorinek@earthlink.net.

Sincerely,

Lauri L. Korinek, M.A. (Ph.D. Candidate)
Principal Investigator

APPENDIX F

Second Email Advertisement

ATTENTION:

Physicians Ages 50 to 80

You are invited to participate in a research project. You will have the opportunity to take a computer-based neuropsychological screen designed specifically for physicians.

The purpose of the study is to collect a normative sample from successfully practicing physicians.

In appreciation for your participation you will

*** Receive feedback about your testing**

*** Receive financial compensation or a donation will be made to your favorite charity**

*** Receive food - a boxed meal**

We have had a good response from physicians who are 30 to 50 years old, but we still need normative data on physicians 50 to 80 years old. If you are interested in the study, please contact

Lauri L. Korinek at 303-324-0451.

Or Read On For More Information

Dear (*Medical Institution)* Physician,

You are invited to participate in a study to collect normative data on a screen of neuropsychological functioning that was designed specifically for physicians. The MicroCog (Powell et al., 1993), a computer-based assessment, was originally developed to assess cognitive functioning of physicians. However, due to changes in the original test, no current normative sample of physicians exists. In addition to its previous uses, we hope that this test will be used to identify and encourage physicians to work beyond the normal retirement age. This will be important in the next few decades as the large number of "baby-boomer" physicians retire.

A second purpose of this study is to gather data from a group of physicians who are successfully practicing medicine and are 30 to 80 years old. We intend to use this information to better understand the cognitive picture of physicians who are practicing successfully in comparison with those who have been involved in competency evaluations.

Participation in this study will involve a one-hour testing session with a research assistant or this investigator. You will receive general feedback about your test results, with the option of receiving more specific feedback upon request. In appreciation for your participation in the study, you will either receive $50 or a donation of $50 will be made to a charity of your choice. Food will also be provided.

All information gathered for this study is confidential. An identification number will be used on all paperwork. Only the researcher will have the list that matches this number with your name, and this list will be kept in an undisclosed secure setting. In addition, when the researcher reports information, it will be reported for the entire group of participants. Individual results will only be reported to the respective participant. Once individual feedback is provided your name will be removed from the list matching your identification number and your name. Upon completion of data collection and feedback, all identifying data will be destroyed.

I hope you will consider participating in this valuable study. Please respond to this e-mail (via telephone or e-mail) regardless of whether you are interested in participating in this study or not, to eliminate future requests for participation. If you are interested in participating in this study, please include a phone number with your response and you will receive a call. Further information

about the study will be provided, and an appointment will be scheduled if you choose to participate. If you have further questions, please contact Lauri L. Korinek at 303-324-0451 or at LLKorinek@earthlink.net.

Sincerely,

Lauri L. Korinek, M.A. (Ph.D. Candidate)
Principal Investigator

APPENDIX G

Invitation Letter to CPEP Physicians

Dear Dr. XXXXX,

You are invited to participate in a study to collect normative data on a screen of neuropsychological functioning that was designed specifically for physicians. The MicroCog (Powell et al., 1993) was originally developed to assess cognitive functioning of physicians. However, due to changes in the original test, no current normative sample of physicians exists. In addition to its previous uses, we hope that this test will be used to identify and encourage physicians to work beyond the normal retirement age. This will be important in the next few decades as the large number of "baby-boomer" physicians retire.

A second purpose of this study is to gather data from a group of physicians (sorry, no residents) who are successfully practicing medicine and are 30 to 80 years old. We intend to then use this information to better understand the cognitive picture of physicians who are practicing successfully in comparison with those who have been involved in competency evaluations.

Participation in this study will involve a one-hour testing session with a research assistant or this investigator. You will receive general feedback about your test results, with the option of receiving more specific feedback upon request. In appreciation for your participation in the study, you will either receive $50 or a donation of $50 will be made to a charity of your choice. Food will also be provided.

All information gathered for this study is confidential. An identification number will be used on all paperwork. Only the researcher will have the list that matches this number with your name, and this list will be kept in an undisclosed secure setting. In addition, when the researcher reports information, it will be reported for the entire group of participants. Individual results will only be reported to the respective participant. Once individual feedback is provided your name will be removed from the list matching your identification number and your name. Upon completion of data collection and feedback, all identifying data will be destroyed.

I hope you will consider participating in this valuable study. Enclosed, you will find a participant response card. Please fill out and return the card in the envelope provided, regardless of whether you are interested in participating in this study or not, to eliminate future requests for participation. If you indicate interest, you will receive a call. Further information about the study will be

provided, and an appointment will be scheduled if you choose to participate. If you have further questions, please contact Lauri L. Korinek at 303-324-0451 or at LLKorinek@earthlink.net.

Sincerely,

Lauri L. Korinek, M.A. (Ph.D. Candidate)
Principal Investigator

APPENDIX H

Participation Form

Participation Form

Please indicate your level of interest in participating in this study. Also, include a phone number where you can be reached in order to provide you with more information or to arrange an appointment. If you are not interested in participating in the study, please return the card anyway, so you will not be bothered with future requests for participation. Complete this card and return it in the stamped envelope provided.

_____I am interested in your study. Please contact me to provide further information.

Phone Number_____

_____I am interested in your study. Please contact me to set up an appointment.

Phone Number_____

_____I am not interested in your study. Please do not contact me.

Thank you for your time.

Lauri L. Korinek, M.A.
Principal Investigator

APPENDIX I

CPEP Support Letter

CPEP
Center for *Personalized*
Education for Physicians

«Fname» «LName», «Degree»
«Home_Address»
«City», «State» «Zip»

January 4, 2005

Dear «Greeting» «LName»:

As a CPEP clinical consultant, you are a valued and respected member of the Colorado medical community. Since you are a leader within your specialty, you would be an ideal participant for a study that has been reviewed and approved by CPEP. I hope you will take the time to read the enclosed information and consider participating.

Lauri Korinek, a doctoral candidate at DU, is recruiting participants to be part of a control group for a study on the neuropsychological characteristics of physicians. Participants must be in good standing in the medical community and have strong clinical skills. If you choose to participate, you would be part of an important clinical study that will improve our understanding of the cognitive function of physicians.

While CPEP supports this research project, we will not be informed of your participation in the study unless you choose to be tested at the CPEP testing site. If you choose to take the exam at CPEP, you should be aware that your confidentiality cannot be protected completely. It is possible that staff members might see you and therefore know that you participated in the study. However, your test results will remain completely confidential and no one at CPEP will have access to this information. If you are concerned about this possibility, you are welcome to take the exam at one of the other testing sites.

CPEP deeply appreciates your work on behalf of the medical community and continued interest in helping with CPEP's competency assessment services. Please feel free to call Ms. Korinek for more information on this study.

Sincerely,

Elizabeth Grace, M.D.
Associate Medical Director

Note: Lauri Korinek is the sister-in-law of CPEP's Executive Director, Beth Korinek. Lauri Korinek is a doctoral candidate in counseling psychology at DU. She is working closely with Laetitia Thompson, Ph.D. Dr. Thompson is an Associate Professor in Psychiatry and Neurology at the University of Colorado Health Sciences Center and is a CPEP consultant.

A National Leader in Evaluating and Enhancing Physician Performance

14001 E. Iliff Avenue, Suite 206 • Aurora, Colorado 80014 • Phone: (303) 750-7150 • Fax: (303) 750-7171 • Web: www.cpepdoc.org

APPENDIX J

University of Denver - Consent Form

Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations.

You are invited to participate in a study of the neuropsychological characteristics of physicians. The study is being conducted by Lauri L. Korinek, M.A. in order to fulfill the requirements for a doctorate in Counseling Psychology at the University of Denver. The results of this research study will be used to learn more about the neuropsychological characteristics of a normative sample of physicians. The project is being supervised by Cyndy McRae, Ph.D., University of Denver, Denver, CO 80208, 303-871-2475. If you have any questions regarding this study please contact Lauri L. Korinek, M.A., 303-324-0451 or Cyndy McRae, Ph.D.

The study will take about 60 minutes to complete. Participation will involve responding to a computerized neuropsychological screen developed for physicians. This test is only a screen and is not considered a definitive clinical evaluation. Your involvement is completely voluntary. You may choose to not participate at any time during the test and are free to withdraw from the study at any time. Refusal to answer a question or withdrawal from participation involves no penalty.

The researcher will treat all information gathered for this study as confidential. This means that only the researcher will have access to the information you provide. An identification number will be used on all paperwork. Only the researcher will have the list that matches this number with your name, and this list will be kept in a secure setting. In addition, when the researcher reports information, it will be reported for the entire group of participants. Individual results will only be reported to the respective participant. Once individual feedback is provided your name will be removed from the list matching your identification number and your name, within 30-days after your test date. Upon completion of data collection and feedback, all identifying data will be destroyed.

Although this research does not address the following, I am required to inform you that there are two exceptions to the promise of confidentiality. Any information you reveal concerning suicide, homicide, or child abuse and neglect is required by law to be reported to the proper authorities. In addition, should any information contained in this study be the subject of a court order, the University of Denver might not be able to avoid compliance with the order or subpoena.

The benefits of being involved in this study include the ability to contribute to improved services for physicians and information about your cognitive performance on an assessment designed for physicians. If your performance on the neuropsychological screen is in the expected range, you will receive a letter informing you of this. If you would like a copy of the results of the study, the researcher will be happy to provide one for you. The results of the study will be reported as group data only. You will receive $50 in appreciation for you participation in the study. Food will also be provided. If you should decide to not complete the testing, compensation will still be provided. Although potential risks are minimal in this study, it is possible that

neuropsychological testing may be upsetting for some people. It is also possible that some neuropsychological problems may be suggested by the screen. If this occurs, this researcher will contact you by phone to inform you of the results. You will be the only person to receive this information. This researcher will then provide the telephone number of the Colorado Physicians Health Program for assistance in obtaining a confidential evaluation.

If you have any concerns or complaints about how you were treated during the research sessions, please contact Dr. Stephen Shirk, Acting Chair, Institutional Review Board for the Protection of Human Subjects, at (303) 871-2484, or Dawn Nowak, Office of Sponsored Programs at (303) 871-4052, or write to either at the University of Denver, Office of Sponsored Programs, 2199 S. University Blvd., Denver, CO 80208.

Please read the following statement and sign below if you agree to participate in this study.
I have read and understood the foregoing description of the study called: Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations. I have asked for and received a satisfactory explanation of any language that I did not fully understand. I agree to participate in this study, and I understand that I may withdraw my consent at any time without penalty. I have received a copy of the consent form.

_____ _____

Signature Date

APPENDIX K

Colorado Multiple Institutional Review Board - Consent Form

Consent Form Approval

Date: 10/06/2004 Valid For Use Through: 10/06/2005

COLORADO MULTIPLE INSTITUTIONAL REVIEW BOARD

Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations.

Principal Investigator: Lauri L. Korinek, M.A., (Ph.D. Candidate)

SUBJECT CONSENT FORM

01/14/2005 Version Three/COMIRB Protocol Number 04-0610

Project Description

You are being asked to take part in a research study comparing the neuropsychological characteristics of two physician groups, physicians who completed a competency evaluation and a control group of physicians. You are being asked to participate in this study as a member of the control group because the computer-based neuropsychological instrument, The MicroCog (Powell et al., 1993), was originally developed to assess cognitive functioning of physicians. However, due to changes in the original test, no current normative sample of physicians exists. Therefore one purpose of this research is to collect normative data on the MicroCog. In addition to its previous uses, we hope that this test will be used to identify and encourage physicians to work beyond the normal retirement age. This will be important in the next few decades as the large number of "baby-boomer" physicians retire.

A second purpose of this study is to gather data from a group of physicians who are successfully practicing medicine and are 30 to 80 years old. We intend to then use this information to better understand the cognitive picture of physicians who are practicing successfully in comparison with those who have been involved in competency evaluations.

This form is designed to provide information about this study so that you can make an informed decision about whether you would like to participate or not. Participation in this study is entirely voluntary.

Up to 75 local participants will be enrolled in this research study.

Procedures

If you agree to participate, you will be asked to do the following things: First you will be asked to read and complete this informed consent form, the Authorization B: Enrollment into Research HIPPA form, a demographic questionnaire, and a

Physician Information Sheet. Next, the examiner will provide an explanation of the testing procedures and familiarize you with the appropriate keys and procedures for taking the test. You will then complete the MicroCog, which will be administered on a laptop computer with an external keyboard attached. You will sit at a desk or table with the computer in front of you to complete the test. The examiner will remain in the testing room for the entire administration to answer any questions you might have.
The MicroCog takes 45 to 60 minutes to complete.

After you complete the assessment, the examiner will save the test results on the computer under the appropriate number code. The test results will be reviewed within a few days by this researcher. The examiner will not view the test results; only the principle investigator will have access to your identified individual results. You will be given $50 in appreciation for your time, or a donation of $50 will be made to a local charity. A boxed meal will be provided after the testing session.

The examiner will offer to provide the results of the entire study (group data from both groups) to you. If you want the results of group data, please mark column on the Physician Information Sheet. Once the study is completed, group results will be sent to all requesting physicians. Once this is completed, the testing session will end. The entire session should take about 60 – 75 minutes to complete.
If you perform within the expected range, this researcher will mail a form letter within two weeks of the testing session. If you would like more specific feedback, this researcher will set up a time for an individual meeting. If you perform outside the expected range, this researcher will contact you, personally, by telephone and will provide feedback, along with the recommendation that you contact the Colorado Physicians Health Program (CPHP) for assistance. Through the CPHP you can obtain confidential neuropsychological testing, rehabilitation, and counseling.
The testing will be conducted by a person who obtained her Ph.D. in Counseling Psychology and this investigator. Before testing, the two examiners will receive training in administering the assessment and in maintaining confidentiality. This investigator will supervise the research assistant. Dr. McRae (University of Denver) and Dr. Thompson (University of Colorado Health Sciences Center) will supervise this investigator.

Discomforts and Risks
Although potential risks are minimal in this study, it is possible that neuropsychological testing may be upsetting for some people. It is also possible that the screen may suggest some neuropsychological problems. If this occurs,

this researcher will contact you by phone to inform you of the results. You will be the only person to receive this information. This researcher will then provide the telephone number of the Colorado Physicians Health Program for assistance in obtaining a confidential evaluation.
The study may include risks that are unknown at this time.

Benefits
This study is designed for the researcher to learn more about the neuropsychological characteristics of two physician groups. This study is not designed to treat any illness or to improve your health. The benefits of being involved in this study include the ability to contribute to better informed services for physicians and information about your cognitive performance on an assessment designed for physicians. If you would like a copy of the results of the study, the researcher will be happy to provide one for you. A boxed meal will be provided. If you should decide to not complete the testing, compensation will still be provided.

Study Sponsor
There is no external funding for this study. The principle investigator is a Ph.D. candidate at the University of Denver, Department of Counseling Psychology. If you have any concerns or complaints about how you were treated during the research sessions, please contact Dr. Stephen Shirk, Acting Chair, Institutional Review Board for the Protection of Human Subjects, at (303) 871-3306, or Dawn Nowak, Office of Sponsored Programs at (303) 871-4052, or write to either at the University of Denver, Office of Sponsored Programs, 2199 S. University Blvd., Denver, CO 80208.

Cost to Subject
There is no cost to you for participating in this study. There will be no charge for any procedures required by the study.

Subject Payment
You will be paid up to $50.00, or a $50.00 donation will be made to a charity of your choice, in appreciation for your participation in this study.

Study Withdrawal
Taking part in this study is voluntary. You have the right to choose not to take part in this study.

If you choose to take part, you have the right to stop at any time. If there are any new findings during the study that may affect whether you want to continue to take part, you will be told about them.

The study investigator may decide to stop your participation without your permission, if he or she thinks that being in the study may cause you harm, or for any other reason.

Invitation for Questions

The researcher carrying out this study is Lauri L. Korinek, M.A. (Ph.D. Candidate). You may ask any questions you have now. If you have questions later, you may call Lauri Korinek at 303-324-0451. You will be given a copy of this form to keep.

If you have questions regarding your rights as a research participant, please call the Colorado Multiple Institutional Review Board (COMIRB) office at (303) 724-1055. Denver Health Medical Center is a participating site for this study.

If you have any concerns or complaints about how you were treated during the research sessions, please contact Dr. Stephen Shirk, Chair, Institutional Review Board for the Protection of Human Subjects, at (303) 871-3306, or Dawn Nowak, Office of Sponsored Programs at (303) 871-4052, or write to either at the University of Denver, Office of Sponsored Programs, 2199 S. University Blvd., Denver, CO 80208.

Confidentiality

We will make every effort to keep your research records confidential, but it cannot be guaranteed. Records that identify you (including your medical records) and the consent form signed by you, may be looked at by a regulatory agency such as:

- Federal agencies that oversee human subject research
- Colorado Multiple Institution Review Board

The results of this research may be presented at meetings or in published articles only in group data format. Your name will be kept private. An identification number will be used on all paperwork. Only the researcher will have access to the list that matches this number with your name and your individual test results. The list and your individual results will be kept in two different undisclosed secures settings. Individual results will only be reported to the respective participant. Once individual feedback is provided your name will be removed from the list matching your identification number and your name within 30 days of your test date. Upon completion of data collection and feedback, all identifying data will be destroyed. You will also be asked to sign a separate

authorization form. This form will explain who will have access to your protected health information.

AUTHORIZATION:
I have read this paper about the study or it was read to me. I understand the possible risk and benefits of this study. I know that being in this study is voluntary. I choose to be in this study. I know I can stop being in this study. I will get a copy of this consent form. (Initial all the previous pages of the consent form).

Signature:_____
Participant

Print Name _____**Date**_____

Consent form explained by: _____
Examiner

Print Name _____ **Date**_____

Investigator - Lauri L. Korinek (Ph.D. Candidate) _____
Date_____

APPENDIX L

University of Denver

Consent Form for Participants Tested at CPEP Test Site

Informed Consent For Participants Tested at the Center for Personalized Education for Physicians

Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations.

You are invited to participate in a study of the neuropsychological characteristics of physicians. The study is being conducted by Lauri L. Korinek, M.A. in order to fulfill the requirements for a doctorate in Counseling Psychology at the University of Denver. The results of this research study will be used to learn more about the neuropsychological characteristics of a normative sample of physicians. The project is being supervised by Cyndy McRae, Ph.D., University of Denver, Denver, CO 80208, 303-871-2475. If you have any questions regarding this study please contact Lauri L. Korinek, M.A., 303-324-0451 or Cyndy McRae, Ph.D.

The study will take about 60 minutes to complete. Participation will involve responding to a computerized neuropsychological screen developed for physicians. This test is only a screen and is not considered a definitive clinical evaluation. Your involvement is completely voluntary. You may choose to not participate at any time during the test and are free to withdraw from the study at any time. Refusal to answer a question or withdrawal from participation involves no penalty.

The researcher will treat all information gathered for this study as confidential. This means that only the researcher will have access to the information you provide. An identification number will be used on all paperwork. Only the researcher will have the list that matches this number with your name, and this list will be kept in a secure setting. In addition, when the researcher reports information, it will be reported for the entire group of participants. Individual results will only be reported to the respective participant. Once individual feedback is provided your name will be removed from the list matching your identification number and your name, within 30-days after your test date. Upon completion of data collection and feedback, all identifying data will be destroyed.

Although this research does not address the following, I am required to inform you that there are two exceptions to the promise of confidentiality. Any information you reveal concerning suicide, homicide, or child abuse and neglect is required by law to be reported to the proper authorities. In addition, should any information contained in this study be the subject of a court order, the University of Denver might not be able to avoid compliance with the order or subpoena.

Since you chose to take the exam at the Center for Personalized Education for Physicians (CPEP), you should be aware that your confidentiality cannot be protected completely. It is possible that staff members might see you and therefore know that you participated in the study. However, your test results will remain completely confidential and no one at CPEP will have access to this information.

The benefits of being involved in this study include the ability to contribute to improved services for physicians and information about your cognitive performance on an assessment designed for physicians. If your performance on the neuropsychological screen is in the expected range, you will receive a letter informing you of this. If you would like a copy of the results of the study, the researcher will be happy to provide one for you. The results of the study will be reported as group data only. You will receive $50 in appreciation for you participation in the study. Food will also be provided.

If you should decide to not complete the testing, compensation will still be provided. Although potential risks are minimal in this study, it is possible that neuropsychological testing may be upsetting for some people. It is also possible that some neuropsychological problems may be suggested by the screen. If this occurs, this researcher will contact you by phone to inform you of the results. You will be the only person to receive this information. This researcher will then provide the telephone number of the Colorado Physicians Health Program for assistance in obtaining a confidential evaluation.

If you have any concerns or complaints about how you were treated during the research sessions, please contact Dr. Stephen Shirk, Acting Chair, Institutional Review Board for the Protection of Human Subjects, at (303) 871-2484, or Dawn Nowak, Office of Sponsored Programs at (303) 871-4052, or write to either at the University of Denver, Office of Sponsored Programs, 2199 S. University Blvd., Denver, CO 80208.

Please read the following statement and sign below if you agree to participate in this study.

I have read and understood the foregoing description of the study called: Neuropsychological Differences Between a Control Group of Physicians and Physicians Referred for Competency Evaluations. I have asked for and received a satisfactory explanation of any language that I did not fully understand. I agree to participate in this study, and I understand that I may withdraw my consent at any time without penalty. I have received a copy of the consent form.

_____ _____

Signature Date

APPENDIX M

Physician Information Sheet

Physician Information Form

Please write in the name and address to which you would like your test results mailed.

Name_____

Address_____

City_____State_____Zip Code_____

E-mail_____
Please write in the best phone number at which to reach you.

Telephone Number_____

Circle One

Yes No Please send me the final study results.

Please circle one of the following two options.

I would like to receive $50 cash today.

I would like to donate $50 to the following charity_____
.

If you would like to have a donation made, please circle one of the following two options.

Please make the donation in my name.

Please make an anonymous donation to the charity.

APPENDIX N

Feedback Letter

From: Lauri L. Korinek, M.A.
6059 South Quebec Street, Suite 201
Englewood, Colorado 80111
Phone: 303-324-0451

Date:

To: Dr. XXXXXX
Address
City, State Zip Code

Dear Dr. XXXXXX,

Your performance on the MicroCog: Assessment of Cognitive Functioning was within the expected range. Please remember that the MicroCog is a screen and is not intended to be used as a definitive clinical evaluation.

If you would like to receive more specific feedback, please contact this investigator within one week. You will then be contacted for an individual feedback session.

Just as a reminder, your name will be removed from the list that enables this researcher to connect your results with your name 30 days after you completed your testing. If you want more specific individual feedback, a timely response to this correspondence is recommended since I will be unable to identify your test results after 30 days from your testing date.

Thank you for participating in this research project. Your contribution is most valued.

Sincerely,

Lauri L. Korinek, M.A., (Ph.D. Candidate)

APPENDIX O

Feedback Telephone Script

Telephone Script for Physicians Performing Outside the Expected Range on the MicroCog

Researcher: Hello Dr. XXXXX, this is Lauri Korinek, I am the person conducting the research on the MicroCog, the computer-based assessment you recently took. First, I would like to thank you for taking your valuable time to participant in this study. I would like to provide you with feedback on your results from the MicroCog. Is this a good time to talk with you?

Dr. XXXXX: Yes. (If this is not a good time to speak with the physician, this researcher will setup another time to provide feedback.)

Researcher: Before providing your individual feedback, I would like to emphasize that the MicroCog is a screen and is not intended to be used as a definitive clinical evaluation. Performances outside the expected range could be due to a variety of reasons and do not necessarily indicate that there is a cognitive problem. With that in mind, here is your feedback. *Provide feedback for individual.* Do you have any questions about the feedback?

Dr. XXXXX: No. (Yes, then answer the questions.)

Researcher: I would like to provide you with the telephone number of the Colorado Physicians Health Program where you can receive assistance in obtaining a confidential neuropsychological evaluation. That number is 303-860-0122. I would also like to inform you that after this conversation your name will be removed from the list that connects your name with your test results. Therefore, no one will be able to place your name with your test results. Do you have any other questions I can help with?

Dr. XXXXX: No. (Yes, then answer the questions)

Researcher: Thank you again for participating in this study. If I can be of any further assistance, please contact me at 303-324-0451.

APPENDIX P

Permissions

Author Permission to Adapt Table 1

MEMO

To: Lauri Korinek

From: Laetitia Thompson, Ph.D.

Date: March 25, 2005

Re: permission to use table

You have my permission to use Table 18.5 "PACE Neuropsychological Test Battery" from my chapter, Neuropsychological Assessment of Physicians Whose Competency to Practice is being Questioned, in Clinical Neuropsychology and Cost Outcome Research edited by G.P. Prigatano and N.H. Pliskin, published by Psychology Press in New York in your dissertation manuscript.

Publisher Permission to Adapt Table 1

Taylor & Francis
Taylor & Francis Group

Permission Agreement R4549

April 19, 2005

Lauri L. Korinek
University of Denver
6059 South Quebec Street, Suite 201
Englewood, CO
80111

Dear Sir or Madam:

We are pleased to grant permission to you to reprint from the following Taylor and Francis title(s):

1841690252 CLINICIAL NEUROPSYCHOLOGY & COST OUTCOMES RESEARCH/1
No Fee Doctoral Dissertation. Pp. 373 - 392, Table 8.5 on p. 388

To appear within:

Publication: **Neuropsychological Differences Between Physicians**

Publisher: **University of Denver**

Publication Date: **June, 2005**

Territory: **North America**

Language: **English**

Format: **Print & Audio**

Fee: **0.00 USD / DATE DUE July 8, 2005** (INVOICE ATTACHED)

Terms & Conditions:

This permission is a non-transferable grant as described below.

1. Payment is nonrefundable and due in USD. If payment is not received by the DUE DATE, this permission will become void (tax ID # is 13-389-6226). Please send payment payable to:

   Taylor & Francis Books, Inc.
   Attn: Subsidiary Rights & Permissions
   270 Madison Avenue
   New York, NY 10016

2. Permission is granted on a one-time, non-exclusive basis.

3. Each copy and or electronic transmission containing our material must bear the following credit line:

   Copyright (Insert © Year) From (Insert Title) by (Insert Author/Editor Name). Reproduced by permission of Routledge/Taylor & Francis Books, Inc.

4. Upon request, one copy of the published Work must be sent to the Permissions Dept. at Taylor & Francis Books, Inc. free of charge.

Author Permission to Adapt Table 2

llkorinek

From: "John Turnbull" <turnbull@mcmaster.ca>
To: "llkorinek" <llkorinek@earthlink.net>
Sent: Sunday, March 06, 2005 12:48 PM
Subject: Re: Cognitive Difficulty In Physicians

Hello Lauri
Of course you can use a table outlining the tests we used. Unfortunately, I have become aware that there were errors introduced when Table 1 was reformatted by the editor. The concluding tables are correct, as are all the general comments, etc. We are preparing a follow-up paper to include remediation results (vs cognitive impairment), and hopefully we can rectify the table at that point. I would be delighted to hear what you are doing.
JT

llkorinek wrote:

> Dr. Turnbull,
>
> I am writing this email to seek your permission to use a table you included in your article, Cognitive Difficulty in Physicians (2000). I am a Counseling Psychology Doctoral Candidate at the University of Denver. For my dissertation, I am conducting research on the neuropsychological differences of two physician groups, a control group and a group of physicians referred for competency evaluations. In my literature review, I cite your article and adapt a table from that = article. The table provides a list of the tests you used to assess different neuropsychological domains of physicians. The table I am using is not an exact copy of your table. The table I am using only lists the tests you used, without mentioning the domains they assessed.
>
> I would be happy to provide more information about this project if you would like. I appreciate the research you've done on physician cognitive concerns.
>
> Lauri L. Korinek, MA, LPC
> 6059 South Quebec Street, Suite 201
> Englewood, Colorado 80111
> Phone: 303-804-5669
> Fax: 303-814-1109
> LLKorinek@earthlink.net
>
>

3/6/2005

Publisher Permission for Adaptation of Table 2

03/22/2005 11:41 FAX 410 528 4549 LIPPINCOTT 001

FAX: 303-814-1109

LIPPINCOTT
WILLIAMS & WILKINS

03/18/05

LAURI KORINEK
6059 S QUEBEC ST STE 201
ENGLEWOOD, CO 80111-4523

Invoice # B45065889 Customer # 000137808630 FEE: 0.00
Re: , ACADEMIC MEDICINE
Spec Mat: ACM 2000;75(2):178 FG. 1
DISSERTATION

INVOICE AND CONDITIONS

Permission is granted for your requested use. Please sign and date this form and return with payment (if applicable) in the enclosed envelope. Please retain a copy for your files. This permission is subject to the following conditions:

1) A credit line will be prominently placed and include: for books - the author(s), title of book, editor, copyright holder, year of publication; for journals - the author(s), title of article, title of journal, volume number, issue number and inclusive pages.

2) The requestor warrants that the material shall not be used in any manner which may be considered derogatory to the title, content, or authors of the material or to LWW.

3) Permission is granted for one time use only as specified in your correspondence. Rights herein do not apply to future reproductions, editions, revisions, or other derivative works.

4) Permission granted is non-exclusive, and is valid throughout the world in the English language only.

5) LWW cannot supply the requestor with the original artwork or a "clean copy."

6) The requestor agrees to secure written permission from the author (for book material only).

7) Permission is valid if the borrowed material is original to a LWW imprint (Lippincott-Raven Publishers, Williams & Wilkins, Lea & Febiger, Harwal, Igaku-Shoin, Rapid Science, Little Brown and Company, Harper & Row Medical American Journal of Nursing Co, and Urban & Schwarzenberg - English Language).

8) Payment can be made via credit card (Amex, VISA, Discover and MC) or by check.

Card # N/A Exp Date: N/A

Requestor accepts: Lauri L. Korinek Date: 3/22/05

Wolters Kluwer Health

351 West Camden Street, Baltimore, MD 21201-2436 • Tel: 410-528-4000 • LWW.com

Author Permission to Adapt Figures 1 and 2

533 Club Side Drive
Naples, FL 34110-6018

March 10, 2005

Lauri L. Korinek, MA
6059 South Quebec Street, Suite 201
Englewood, CO 80111

Dear Laurie,

I'm very glad to grant you permission to reproduce the figures on pages 1 and 14 from the MicroCog Manual (1993). I believe that it will be necessary also that Harcourt Assessment also sign off.

Parenthetically, I think your research is quite promising. I hope that you will keep me updated on your progress.

Cordially,

Douglas H. Powell
Clinical Instructor in Psychology
Harvard Medical School
douglas_powell@hms.harvard.edu

Publisher Permission to Adapt Figures 1 and 2

Subj: **Permission Request**
Date: 3/9/2005 1:05:13 PM Mountain Standard Time
From: christine_doebbler@harcourt.com
To: LLKorinek@aol.com

Ms. Lauri Korinek
Graduate Student
University of Denver
6059 South Quebec Street
Suite 201
Englewood, CO 80111

Dear Ms. Korinek:

Thank you for your email of March 1 regarding permission to include materials from the MicroCog: Assessment of Cognitive Functioning manual in your dissertation.

Permission is hereby granted for the inclusion of the following Figures in your dissertation:

1. Figure 1.1 Levels of MicroCog Index Scores
2. Figure 1.2 Subtests Contributing to Level 1 Index Scores and Standard and Short Forms

Please ensure that the fully copyright notice from the manual appears with each table, preceded by the words "Reproduced with permission of Publisher, Harcourt Assessment, Inc." .

Thank you for your interest. If you have other questions or needs, please let us know.

Regards,
Christine Doebbler, Director
Legal Services
Harcourt Assessment, Inc.
19500 Bulverde Road
San Antonio, Texas 78259
U.S.A.
(210) 339-5577 phone
(210) 339-5059 fax

NEUROPSYCHOLOGICAL DIFFERENCES BETWEEN PHYSICIANS REFERRED FOR COMPETENCY EVALUATIONS AND A CONTROL GROUP OF PHYSICIANS

An Abstract of a Dissertation

Presented to

The College of Education

University of Denver

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

---

by

Lauri L. Korinek

June 2005

Advisor: Cynthia McRae, Ph.D.

Research conducted through physician evaluation programs suggests that there are neuropsychological concerns among the physicians evaluated (Turnbull et al., 2000; Williams, Williams, & Norcross, 2002). However, more research with physician control groups is needed (Perry & Crean, 2005, Thompson, 2003). The purpose of this study was to compare the neuropsychological differences between a group of physicians referred for competency evaluations and a control group of physicians.

Using the MicroCog (Powell et al., 1993), a computerized neuropsychological screen originally designed for physicians, this study compared the cognitive performance of 267 physicians referred for competency evaluations with a control group of 68 volunteer physicians. During a seven-year period, 1997-2004, physicians referred for competency evaluations, took the MicroCog as a part of their evaluation at the Center for Personalized Education for Physicians. The control group was comprised of practicing physicians whose competency was not in question.

The competency evaluation group had a greater proportion of cognitively impaired individuals and performed significantly lower on scores of processing speed, processing accuracy, and cognitive proficiency than the control group. The control group of physicians performed significantly better than the age- and education- corrected normative sample.

Because there were significant neuropsychological differences between physicians referred for competency evaluations and physicians whose competency was not in question, it appears to be important that neuropsychological assessment be included as part of physicians' competency evaluations. More research is needed to better understand how cognitive impairment impacts physician performance, and physician norms for the MicroCog (Powell et al., 1993) might provide a better comparison group for physicians.